ADARAMOLA, BUKOLA O., M.A. Circulant Matrices On Global Data Analysis. (2017)
Directed by Dr. Haimeng Zhang. 66 pp.

In this thesis, we investigate the unbiasedness of the commonly used covariance and variogram estimators when the underlying process is assumed to be stationary on the circle or axially symmetric on the sphere. We show that the covariance estimator is generally biased while the variogram estimator is unbiased. We represent the variogram estimator as a quadratic form of the observed gridded data whose variance-covariance structure is associated with circulant or block circulant matrices. We then use the spectral decomposition of circulant and block circulant matrices to decompose the variogram estimator as a linear combination of uncorrelated random variates. In particular, if the underlying process on the circle is stationary and Gaussian, we express the variogram estimator as a linear combination of independent and identically distributed χ_1^2 random variates. The same result is also obtained if the process is assumed to be a longitudinally reversible Gaussian process on the sphere.

CIRCULANT MATRICES ON GLOBAL DATA ANALYSIS

by

Bukola O. Adaramola

A Thesis Submitted to
the Faculty of The Graduate School at
The University of North Carolina at Greensboro
in Partial Fulfillment
of the Requirements for the Degree
Master of Arts

Greensboro
2017

Approved by

Committee Chair

I will give special thanks to God and dedicate this work to His greatness. I dedicate my Thesis work to my family. A special gratitude to my loving wife, Ogechi Adaramola for her words of encouragement. I also dedicate this thesis to my children for their support. I will always appreciate all they have done.

APPROVAL PAGE

This thesis written by Bukola O. Adaramola has been approved by the following committee of the Faculty of The Graduate School at The University of North Carolina at Greensboro.

Committee Chair ______________________
Haimeng Zhang

Committee Members ______________________
Sat Gupta

____
Thomas Lewis

____
Scott Richter

____
Dan Yasaki

____
Date of Acceptance by Committee

____
Date of Final Oral Examination

ACKNOWLEDGMENTS

I would like to thank my thesis advisor Dr. Haimeng Zhang at University of North Carolina at Greensboro for his continuous support, patience, motivation and immense knowledge. He consistently allowed this paper to be my own work, but steered me in the right direction whenever he thought I needed it. I could not have a better advisor and mentor.

I will also like to thank all the committee members for their time.....

Finally, I must express my very profound gratitude to my parents, Mr Kayode Adaramola and Mrs Iyabo Adaramola and also to my wife, Mrs Ogechi Adaramola for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of writing this thesis. This accomplishment would not have been possible without them. Thank you all.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

In this chapter we give a brief introduction to some of the basic concepts in spatial statistics. Specifically, we discuss stationarity and intrinsic stationarity, covariance and variogram functions and their properties, complex random variables and vectors, and some properties related to quadratic forms. Literature reviews and the outline of this thesis are presented at the end.

1.1 Random Processes

A *random process* is a collection of random variables $\{X(t) : t \in T\}$ defined in the same probability space. Here T is a parameter set. Usually, T is called the time set. If T is discrete, $\{X(t) : t \in T\}$ is called a *discrete-time process*. Usually, $T = \{0, 1, \ldots\}$. If T is an interval, $\{X(t) : t \in T\}$ is called a *continuous-time process*. Usually, $T = [0, 1]$ or $T = \mathbb{R}^d$, a d-dimensional Euclidean space.

For each element $t \in T$, the *(cumulative) distribution function* of the random function $X(t)$ is defined as

$$F_t(x) = P(X(t) \leq x). \tag{1.1}$$

If $t_1, t_2 \in T$, $(X(t_1), X(t_2))$ represents a bivariate random variable, where its joint distribution function is expressed as

$$F_{t_1,t_2}(x_1, x_2) = P(X(t_1) \leq x_1, X(t_2) \leq x_2). \tag{1.2}$$

In general, for a random process $\{X(t) : t \in T\}$, the *distribution function* for an n-variate random variable $(X(t_1), X(t_2), \cdots, X(t_n))$ is defined as

$$F_{t_1,t_2,\cdots,t_n}(x_1, x_2, \cdots, x_n) = P(X(t_1) \leq x_1, X(t_2) \leq x_2, \cdots, X(t_n) \leq x_n). \tag{1.3}$$

The finite-dimensional distribution functions defined by (1.3) satisfy the following conditions for any permutation $(j_1, j_2, \cdots, j_n)$ of $(1, 2, \cdots, n)$:

(a) The symmetry condition:

$$F_{t_{j_1},t_{j_2},\cdots,t_{j_n}}(x_{j_1}, x_{j_2}, \cdots, x_{j_n}) = F_{t_1,t_2,\cdots,t_n}(x_1, x_2, \cdots, x_n).$$

(b) The compatibility condition:

$$F_{t_1,t_2,\cdots,t_m,t_{m+1},\cdots,t_n}(x_1, x_2, \cdots, x_m, \infty, \cdots, \infty) = F_{t_1,t_2,\cdots,t_m}(x_1, x_2, \cdots, x_m)$$

for any $t_1, t_2, \ldots, t_m, t_{m+1}, \cdots, t_n$, if $m < n$.

1.2 Stationarity

In Euclidean space, a spatial random process $X(t)$ is a collection of spatial random variables that vary over a continuous subset $D \subseteq \mathbb{R}^d$. For any fixed, finite set of elements $t_1, \cdots, t_n \in D$, the n-variate random variate $(X(t_1), \cdots, X(t_n))$ is a n-dimensional random vector associated with its finite dimensional distribution given by

$$F_{t_1,t_2,\cdots,t_n}(x_1, \cdots, x_n) = P(X(t_1) \leq x_1, \cdots, X(t_n) \leq x_n). \tag{1.4}$$

A spatial random field is called *weakly or second-order stationary* if we have

$$E(X(t)) = \mu, \tag{1.5}$$

$$cov(X(t), X(t+h)) = E(X(t) - \mu)(X(t+h) - \mu) = C(h), \tag{1.6}$$

where μ is a constant, $C(h)$ is the covariance function only depending on the displacement h, and $var(X(t)) < \infty$.

Now we introduce a weaker assumption called *intrinsic stationarity* using the variogram. A spatial random field is intrinsically stationary if the increment process $X(t) - X(t+h), t \in \mathbb{R}^d$ is stationary for all leg vectors $h \in \mathbb{R}^d$. In other words, $E(X(t)) = \mu$ (constant) and $E(X(t) - X(t+h))^2 = var(X(t) - X(t+h))$ does not depend on t, where the semivariogram, originally defined by Matheron (1962), is calculated as

$$\gamma(h) = \frac{1}{2}(E(X(t) - X(t+h)))^2, \quad h \in \mathbb{R}^d. \tag{1.7}$$

For a weakly stationary random process, it is easy to see that $\gamma(h) = C(0) - C(h)$, so a weakly stationary process is intrinsically stationary, but not conversely. For example, if one considers two locations t_1 and t_2 for a one-dimensional Brownian Motion with the covariance function given by $cov(X(t_1), X(t_2)) = \min(t_1, t_2)$, one has

$$\begin{aligned}
var(X(t_1) - X(t_2)) &= cov(X(t_1) - X(t_2), X(t_1) - X(t_2)) \\
&= cov(X(t_1), X(t_1)) - 2cov(X(t_1), X(t_2)) + cov(X(t_2), X(t_2)) \\
&= t_1 - 2\min(t_1, t_2) + t_2 \\
&= \begin{cases} t_2 - t_1, & \text{if } t_1 < t_2 \\ t_1 - t_2, & \text{if } t_2 \leq t_1 \end{cases} \\
&= |t_2 - t_1|,
\end{aligned}$$

which is a function of displacement. Therefore, it is intrinsically stationary.

The intrinsic stationarity is generally more applicable in real data analysis. Therefore, the variogram (defined below) is a more preferred tool in characterizing the dependency on geospatial processes.

1.2.1 Validity of Covariance and Variogram Functions

Consider a random process $\{X(t) : t \in D\}$. Let $t_i, t_j \in D$ be two locations. The covariance function of $X(t)$ at t_i and t_j is defined as $C(t_i, t_j) = cov(X(t_i), X(t_j))$.

The variogram function is defined as $2\gamma(t_i, t_j) = var(X(t_i) - X(t_j))$. A real continuous covariance function on $\mathbb{R}^d$ is valid if and only if it is positive definite where

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j C(t_i, t_j) \geq 0 \tag{1.8}$$

for any integer n, any real numbers $\{a_i : 1 \leq i \leq n\}$, and any locations $\{t_i : i = 1, \cdots, n\} \subset \mathbb{R}^d$. Similarly, a variogram function $2\gamma(h)$ is considered valid if it is conditionally negative definite in the sense that

$$\sum_{i=1}^{m}\sum_{j=1}^{m} a_i a_j 2\gamma(t_i, t_j) \leq 0, \tag{1.9}$$

for any real numbers $\{a_i : i = 1, \cdots, n\}$, where $\sum_{i=1}^{m} a_i = 0$ and any locations $\{t_i : i = 1, \cdots, n\} \subset \mathbb{R}^d$.

1.3 Complex Random Vectors

A complex random variable has two parts, the real and imaginary parts. We define $Z = X + iY$ to be a complex random variable where X and Y are real random variables. The probability density function (PDF) of a complex random variable is defined as the joint PDF of its real and imaginary parts. The complex random variable Z can be treated as a random vector $[X, Y]$.

1.3.1 Moments, Variance and Covariance

The k^{th} *moment* of a real random variable X is defined as $m_k = E[X^k]$. Its *variance* is defined as $var[X] = E[(X - E[X])^2]$ with the following properties.

- $var[X] = E[X^2] - E[X]^2$,
- $var[aX] = a^2 var[X]$.

The *covariance* of random variables X and Y is defined as:

$$cov[X,Y] = E[(X - E[X])(Y - E[Y])] = E[XY] - E[X]E[Y].$$

Also, the *covariance matrix* of a random vector $\underset{\sim}{X}$ is defined as:

$$K_x = E[(\underset{\sim}{X} - E[\underset{\sim}{X}])(\underset{\sim}{X} - E[\underset{\sim}{X}])^T]$$

with the properties:

- K_x is symmetric;
- The elements of $\underset{\sim}{X}$ are uncorrelated if its covariance matrix is diagonal.

1.3.2 Complex-Valued Random Vector

Let $\underset{\sim}{V} = \underset{\sim}{V}_r + i\underset{\sim}{V}_i$ and $\underset{\sim}{U} = \underset{\sim}{U}_r + i\underset{\sim}{U}_i$ be complex-valued random vectors. The expectation is given by: $E[\underset{\sim}{V}] = E[\underset{\sim}{V}_r] + iE[\underset{\sim}{V}_i]$.

The second order statistics are completely characterized either by the four real-valued covariance matrices:

$$\begin{aligned} K_{u_R v_R} &= cov[\underset{\sim}{U}_r, \underset{\sim}{V}_r], \\ K_{u_R v_I} &= cov[\underset{\sim}{U}_r, \underset{\sim}{V}_i], \\ K_{u_I v_R} &= cov[\underset{\sim}{U}_i, \underset{\sim}{V}_r], \\ K_{u_I v_I} &= cov[\underset{\sim}{U}_i, \underset{\sim}{V}_i], \end{aligned}$$

or the two complex-valued covariance matrices

$$\begin{aligned} K_{UV} &= E[(\underset{\sim}{U} - E[\underset{\sim}{U}])(\underset{\sim}{V} - E[\underset{\sim}{V}])^*] \\ J_{UV} &= E[(\underset{\sim}{U} - E[\underset{\sim}{U}])(\underset{\sim}{V} - E[\underset{\sim}{V}])^T]. \end{aligned}$$

Here $\underset{\sim}{a}^*$ represents the complex conjugated transpose of a complex vector $\underset{\sim}{a}$ throughout the rest of this thesis, that is, $\underset{\sim}{a}^* = (\bar{\underset{\sim}{a}})^T$.

K_{UV} is called *covariance matrix*, while J_{UV} is called *pseudo-covariance matrix*. The following relations hold.

$$\begin{aligned} K_{UV} &= K_{u_R v_R} + K_{u_I v_I} + i(K_{u_I v_R} - K_{u_R v_I}), \\ J_{UV} &= K_{u_R v_R} - K_{u_I v_I} + i(K_{u_I v_R} + K_{u_R v_I}), \end{aligned}$$

and

$$\begin{aligned} K_{u_R v_R} &= \frac{1}{2}\Re(K_{UV} + J_{UV}), \\ K_{u_I v_I} &= \frac{1}{2}\Re(K_{UV} - J_{UV}), \\ K_{u_I v_R} &= \frac{1}{2}\Im(K_{UV} + J_{UV}), \\ K_{u_R v_I} &= \frac{1}{2}\Im(-K_{UV} + J_{UV}), \end{aligned}$$

where $\Re$ is the real part and $\Im$ is the imaginary part.

1.3.3 Quadratic Forms

Let $\utilde{X} = (X_1, X_2, \ldots, X_n)^T$ be a random vector of dimension n. If $E(\utilde{X}) = \utilde{\mu}$ and $var(\utilde{X}) = \Sigma$, then $Y = \utilde{X}^T A \utilde{X}$ is called a *quadratic form*. Usually $A_{n\times n}$ is a symmetric matrix but not necessary positive definite.

For the above quadratic form, we have the following properties.

Proposition 1.1. $E(Y) = tr(A\Sigma) + \utilde{\mu}^T A \utilde{\mu}$ and if $\utilde{X}$ further follows a multivariate normal distribution, $var(Y) = 2tr(A\Sigma A\Sigma) + 4\utilde{\mu}^T A\Sigma A\utilde{\mu}$.

Proof. The first result is a simple calculation from Mathematics Statistics:

$$\begin{aligned} E(Y) &= E(tr(\utilde{X}^T A \utilde{X})) = E(tr(A\utilde{X}\utilde{X}^T)) = tr(AE(\utilde{X}\utilde{X}^T)) \\ &= tr(A(var(\utilde{X}) + E(\utilde{X})E(\utilde{X}^T))) \\ &= tr(A\Sigma) + tr(A\utilde{\mu}\utilde{\mu}^T) = tr(A\Sigma) + tr(\utilde{\mu}^T A\utilde{\mu}) = tr(A\Sigma) + \utilde{\mu}^T A\utilde{\mu}. \end{aligned}$$

The proof of the second result is tedious. One can refer to Theorem 1.6 of Seber and Lee (2003) for a more general result.

Corollary 1.1. If $\underset{\sim}{\mu} = \underset{\sim}{0}$, then we have $E(Y) = tr(A\Sigma)$, and if additionally $\underset{\sim}{X}$ further follows a multivariate normal distribution, $var(Y) = 2tr(A\Sigma A\Sigma)$.

1.4 Literature Reviews

In spatial statistics, one of the most important questions of interest to geo-scientists is how to model the dependency of a spatial process based on observed data and then use these models to make predictions at unobserved locations. Such a dependency modeling often involves finding the covariance and/or variogram functions. However, it is a high-dimensional problem since in geostatistics normally one data value is available at each location. Hence, no statistical inference would be possible if no further assumptions are given. For example, if data are observed at 10 locations, in order to construct the variance-covariance matrix based on the observed (random) data, one needs to estimate 55 pairwise variance-covariance entries, which is impossible.

There has been an extensive study of covariance and variogram estimators based on the method of moments (MOM) in $\mathbb{R}^d$, see Cressie (1993) for a comprehensive discussion. In recent years, data from global networks and satellite sensors have been used to monitor a wide array of processes and variables, such as temperature, precipitation, etc. In the modeling and analysis of such global-scale data, the covariance and variogram estimators have been extensively used to model the dependency. For example, Stein (2007) used the variogram estimator to model global variations in TOMS (Total Ozone Mapping Spectrometer) measurements. However, it is surprising that no study of their properties is discussed in the literature. Vanlangenberg (2016) considered the asymptotics of these estimators for stationary processes on the circle. Although it has been known that such estimators are asymptotically unbiased and consistent when modeling the stationary process on $\mathbb{R}^d$, his findings on the circle showed the contrast. More explicitly, he found out that the MOM variogram estimator is unbiased but inconsistent under the assumption of Gaussianity. In this thesis, we focus on the unbiasedness and distributional properties of MOM covariance and variogram estimators on both the circle and the sphere based on the spectral decomposition of circulant and block circulant matrices.

1.5 Outline of This Thesis

In Chapter 2, we introduce the Fourier matrix, circulant matrix, and block circulant matrix. We discuss the unitary-diagonalization and real orthogonal - diagonalization of circulant and block circulant matrices. In particular, a closed form for calculating their eigenvalues/diagonal matrices is presented. Simple examples are given to illustrate the concepts and calculations. In Chapter 3, we explore some properties of commonly used covariance and variogram estimators based on MOM on the circle. We first provide a spectral representation of the variance-covariance matrix of an observed gridded random vector on the circle, and then we evaluate the unbiasedness and distributional properties of the covariance and variogram estimators based on MOM. More explicitly, we decompose the variogram estimator as a linear combination of independent χ^2 random variates when the underlying process is assumed to be Gaussian. Similar results are obtained for covariance and variogram estimators based on MOM on the sphere in Chapter 4. Finally, a summary of further research directions is provided in Chapter 5.

CHAPTER II

CIRCULANT MATRIX

In this chapter, we first give some basic properties regarding roots of unity. We then discuss Krylov matrices, including the Fourier matrix as a special case. In Section 2.4, we introduce the circulant matrix as well as its spectral decomposition where the Fourier matrix is its corresponding unitary matrix. Special attention has been given to the real-valued symmetric circulant matrix, where the eignevalues are real and duplicated while the Fourier matrix can be replaced by a real orthogonal matrix. Parallel results are also presented for block-circulant and symmetric block circulant matrices in Section 2.5.

2.1 Roots of Unity

Let $\omega_{(n)} = \exp(2\pi i/n)$. Then $\omega_{(n)}$ is a primitive n^{th} root of unity because it generates all other n^{th} roots of unity. We then let $\omega_{(n)}^j = \exp(2\pi ij/n)$, which has the following properties:

(1) All $\omega_{(n)}^j, j = 0, 1, \ldots, n-1$, are roots of the polynomial $p(x) = x^n - 1$.

(2) $\overline{\omega_{(n)}}^j = \omega_{(n)}^{-j}$.

(3) If p divides n, then

$$\omega_{(n)}^p = \exp(2\pi ip/n) = \exp(2\pi i/(n/p)) = \omega_{(n/p)} = n/p^{\text{th}} \text{ root of unity.}$$

Let $W_n = \mathrm{diag}(\omega_{(n)}^0, \omega_{(n)}^1, \ldots, \omega_{(n)}^{n-1})$. As example, let $\omega_{(4)} = \exp(2\pi i/4) = \exp(\pi i/2) = i$. Then $W_4 = \mathrm{diag}(\omega_{(4)}^0, \omega_{(4)}^1, \omega_{(4)}^2, \omega_{(4)}^3) = \mathrm{diag}(1, i, -1, -i)$.

2.2 Krylov Matrix

Let $A \in \mathbb{C}^{n\times n}$. The *Krylov matrix* of A generated by a vector $\underset{\sim}{b} \in \mathbb{C}^n$ is given by

$$K(A, \underset{\sim}{b}) = [\underset{\sim}{b}, A\underset{\sim}{b}, A^2\underset{\sim}{b}, \ldots, A^{n-1}\underset{\sim}{b}].$$

Now let $W_n = \mathrm{diag}(\omega_{(n)}^0, \omega_{(n)}^1, \ldots, \omega_{(n)}^{n-1})$ and $\underset{\sim}{e} = (1, 1, \ldots, 1)^T$. Hence,

$$F_n = K(W_n, \underset{\sim}{e}) = \begin{bmatrix} 1 & \omega_{(n)}^0 & \ldots & \omega_{(n)}^{n-1} \\ 1 & \omega_{(n)}^1 & \ldots & \omega_{(n)}^{2(n-1)} \\ 1 & \omega_{(n)}^2 & \ldots & \omega_{(n)}^{3(n-1)} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \omega_n^{n-1} & \ldots & \omega_{(n)}^{(n-1)(n-1)} \end{bmatrix}.$$

which implies that $(F_n)_{jk} = (\omega_n)^{(j-1)(k-1)}$ where j is the row number and k is the column number. F_n is called the *Fourier Matrix.* For example, since $\omega_4 = i$, there holds

$$F_4 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{bmatrix}.$$

Properties of F_n

- F_n is symmetric;
- $F_n^* F_n = nI$.

Now $F_n^* F_n = nI$ which implies $\frac{F_n^* F_n}{n} = I$ which implies $\frac{F_n^*}{\sqrt{n}} \frac{F_n}{\sqrt{n}} = I$. Then, $U_n = \frac{F_n}{\sqrt{n}}$ and $U_n^* = \frac{F_n^*}{\sqrt{n}}$ are unitary matrices.

2.3 Companion Matrix

The *companion matrix* of a monic polynomial $p(x) = c_0 + c_1 x + c_2 x^2 + \ldots + c_{n-1} x^{n-1} + x^n$ is the square matrix

$$C := \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ddots & 1 \\ c_0 & c_1 & c_2 & \cdots & c_{n-1} \end{pmatrix}.$$

For $p(x) = x^n + 1$, we have

$$C := \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ddots & 1 \\ 1 & 0 & 0 & \cdots & 0 \end{pmatrix},$$

which is also called the *Downshift Permutation*. Then F_n is a *Vandermonde Matrix* with $W_n F_n = F_n C$ where C is the companion matrix for $p(x) = x^n + 1$. Since $W_n = \mathrm{diag}(\omega_{(n)}^j)$,

$$W_n F_n = F_n C \Rightarrow W_n U_n = U_n C \Rightarrow U_n^* W_n U_n = C.$$

2.4 Circulant Matrix

A *circulant matrix* R_b taken with respect to the vector $\underset{\sim}{b} = (b_0, b_1, \ldots, b_{n-1})^T$ is

$$R_b = \begin{pmatrix} b_0 & b_1 & \cdots & b_{n-1} \\ b_{n-1} & b_0 & \cdots & b_{n-2} \\ \vdots & & \ddots & \vdots \\ b_1 & b_2 & \cdots & b_0 \end{pmatrix}$$

which $implies R_b = K(C, \underset{\sim}{b}) = [\underset{\sim}{b}, C\underset{\sim}{b}, C^2\underset{\sim}{b}, \ldots, C^{n-1}\underset{\sim}{b}]$. Then, for any circulant matrix, $R_b = U_n \mathrm{diag}(F_n \underset{\sim}{b}) U_n^*$.

Proof:

$$\begin{aligned}
U_n^* R_b U_n &= \operatorname{diag}(F_n \underset{\sim}{b}). \\
U_n^* R_b U_n &= U^* K(C, \underset{\sim}{b}) U \\
&= U^*[\underset{\sim}{b}, C\underset{\sim}{b}, C^2\underset{\sim}{b}, \ldots, C^{n-1}\underset{\sim}{b}]U \\
&= [U^*\underset{\sim}{b}, U^*CUU^*\underset{\sim}{b}, \ldots, (U^*CU)^{n-1}U^*\underset{\sim}{b}] \\
&= [U^*\underset{\sim}{b}, \omega U^*\underset{\sim}{b}, \ldots, \omega^{n-1}U^*\underset{\sim}{b}] \\
&= [D\underset{\sim}{e}, \omega D\underset{\sim}{e}, \ldots, \omega^{n-1}D\underset{\sim}{e}] \\
&= D[\underset{\sim}{e}, \omega\underset{\sim}{e}, \ldots, \omega^{n-1}\underset{\sim}{e}],
\end{aligned}$$

where $D = \operatorname{diag}(U^*\underset{\sim}{b})$. Hence,

$$U^*K(C, \underset{\sim}{b}) = \operatorname{diag}(F, \underset{\sim}{b})\frac{F_n}{\sqrt{n}} = \operatorname{diag}(F, \underset{\sim}{b})U^*, \quad \Rightarrow \quad U^*K(C, \underset{\sim}{b})U = \operatorname{diag}(F, \underset{\sim}{b}).$$

2.4.1 Eigenvalues and Eigenvectors of Circulant Matrix

Let $\omega_j = \exp(2\pi(j-1)/n)$. The circulant matrix R_b has n eigenvalues given by

$$\lambda_j \;\; = b_0 + b_1 w_j + \ldots + b_{n-1} w_j^{n-1}, \qquad j = 1, 2, \ldots, n,$$

and has n orthogonal eigenvectors given by

$$\underset{\sim}{w}^{(j)} = \begin{bmatrix} 1 \\ \omega_j \\ \omega_j^2 \\ \vdots \\ \omega_j^{n-1} \end{bmatrix} \tag{2.1}$$

for $j = 1, 2, \ldots, n$. If R_b is a real symmetric circulant matrix, that is, $b_i = b_{n_i}$, then its eigenvalues are real. More specifically, for even $n = 2N$ the eigenvalues $\lambda_j = \lambda_{n-j}$ (there are either two eigenvalues or none with odd multiplicity), for odd $n = 2N - 1$, the eigenvalue λ_1 equal to any λ_j for $1 \leq j \leq N$ or λ_0 occurs with odd multiplicity. A square matrix $\mathbf{A}$ is called *Hermitian* if $A^* = A$. If the matrix is real, then $A^T = A$. Each Hermitian matrix has a full set of orthogonal eigenvectors, each with real eigenvalues.

In addition, if we specify the following real orthogonal matrix,

$$\begin{aligned} Q &= \left(\frac{u^{(0)}}{\sqrt{2N}}, \frac{u^{(1)}}{\sqrt{N}}, \frac{v^{(1)}}{\sqrt{N}}, \frac{u^{(2)}}{\sqrt{N}}, \frac{v^{(2)}}{\sqrt{N}}, \cdots, \frac{u^{(N-1)}}{\sqrt{N}}, \frac{v^{(N-1)}}{\sqrt{N}}, \frac{u^{(N)}}{\sqrt{2N}} \right) \\ &= \frac{1}{\sqrt{N}} \left(\frac{u^{(0)}}{\sqrt{2}}, u^{(1)}, v^{(1)}, u^{(2)}, v^{(2)}, \cdots, u^{(N-1)}, v^{(N-1)}, \frac{u^{(N)}}{\sqrt{2}} \right) \end{aligned} \tag{2.2}$$

with

$$\begin{aligned} u^{(l)} &= (u_j^{(l)}) = (\cos(2jl\pi/2N)), j = 1, 2, \cdots, 2N, l = 0, 1, \cdots, N, \\ v^{(l)} &= (v_j^{(l)}) = (\sin(2jl\pi/2N)), j = 1, 2, \cdots, 2N, l = 1, \cdots, N-1, \end{aligned}$$

then R_b can be diagonalized as

$$R_b = Q\text{diag}(\lambda_j, j = 1, 2, \ldots, n)Q^T.$$

2.5 Block Circulant Matrix

A *block circulant matrix* $A \in \mathbb{R}^{np \times np}$ is of the form given below.

$$\begin{aligned} A &= circ(A_0, A_1, \cdots, A_{n-1}) \\ &= \begin{pmatrix} A_0 & A_1 & A_2 & \cdots & A_{n-2} & A_{n-1} \\ A_{n-1} & A_0 & A_1 & \cdots & A_{n-3} & A_{n-2} \\ A_{n-2} & A_{n-1} & A_0 & \cdots & A_{n-4} & A_{n-3} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ A_1 & A_2 & A_3 & \cdots & A_{n-1} & A_0 \end{pmatrix}, \end{aligned}$$

where each A_k is of dimension $p \times p$.

2.5.1 Eigenvalues and Eigenvectors of Block Circulant Matrix

According to De Mazancourt and Gerlic (1983), we set

$$E_j = \begin{pmatrix} I_p \\ \omega_j I_p \\ \omega_j^2 I_p \\ \vdots \\ \omega_j^{n-1} I_p \end{pmatrix}, \qquad j = 1, 2, \cdots, n,$$

with $\omega_k = (\exp((j-1)2\pi i/n)), j = 1, \cdots, n$ and I_p being an identity matrix of dimension $p \times p$. Let

$$P = \frac{1}{\sqrt{n}}(E_1, E_2, E_3, \cdots, E_n).$$

Then

$$A = P \times \begin{pmatrix} S_1 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & S_2 & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & S_3 & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & S_n \end{pmatrix} \times P^* := P\mathrm{diag}(S_1, S_2, \ldots, S_n)P^*,$$

where S_j is of dimension $p \times p$ with

$$S_j = \sum_{m=0}^{n-1} \omega_j^m A_m.$$

Here the block S_k of dimension $p \times p$ is the discrete Fourier transform of the A_k blocks. It is also important to note that the matrix P which diagonalizes A is completely independent of the blocks of A. This means that P is also the general matrix that transforms a block-diagonal matrix into a block-circulant matrix (BCM). That is, for any block-diagonal matrix D, the product PDP^* is a BCM.

Now we consider a *block-symmetric block circulant matrix* (here block - symmetric of A is in the sense of $A_j = A_{n-j}$ throughout this Thesis). We define the following orthogonal matrix:

$$Q = (Q_1, Q_2, Q_3, \cdots, Q_n)$$

where

$$Q_1 = \frac{1}{\sqrt{n}} \begin{pmatrix} I_p \\ I_p \\ I_p \\ \vdots \\ I_p \end{pmatrix}, Q_2 = \frac{1}{\sqrt{n/2}} \begin{pmatrix} I_p \\ \cos(1\pi/n) I_p \\ \cos(2\pi/n) I_p \\ \cos(3\pi/n) I_p \\ \vdots \\ \cos((n-1)\pi/n) I_p \end{pmatrix},$$

$$Q_3 = \frac{1}{\sqrt{n/2}} \begin{pmatrix} I_p \\ \sin(1\pi/n) I_p \\ \sin(2\pi/n) I_p \\ \sin(3\pi/n) I_p \\ \vdots \\ \sin((n-1)\pi/n) I_p \end{pmatrix}, \ldots,$$

and

$$Q_{n-2} = \frac{1}{\sqrt{n/2}} \begin{pmatrix} I_p \\ \cos(2(N-1)\pi/n)I_p \\ \cos(4(N-1)\pi/n)I_p \\ \cos(6(N-1)\pi/n)I_p \\ \vdots \\ \cos(2(n-1)(N-1)\pi/n)I_p \end{pmatrix},$$

$$Q_{n-1} = \frac{1}{\sqrt{n/2}} \begin{pmatrix} I_p \\ \sin(2(N-1)\pi/n)I_p \\ \sin(4(N-1)\pi/n)I_p \\ \sin(6(N-1)\pi/n)I_p \\ \vdots \\ \sin(2(n-1)(N-1)\pi/n)I_p \end{pmatrix}, Q_n = \frac{1}{\sqrt{n}} \begin{pmatrix} I_p \\ -I_p \\ I_p \\ -I_p \\ \vdots \\ -I_p \end{pmatrix}.$$

The block-symmetric block circulant matrix with symmetric block A_m can be decomposed as

$$\Sigma = Q\text{diag}(S_1^{(\Sigma)}, S_2^{(\Sigma)}, \ldots, S_n^{(\Sigma)})Q^T,$$

where $S_i^{(\Sigma)}, i = 1, 2, \ldots, n$ are $p \times p$ matrices, with

$$S_j^{(\Sigma)} = A_0 + 2\sum_{m=1}^{N-1} \cos(m(j-1)\delta)A_m + \cos((j-1)\delta\pi)A_N, j = 1, 2, \ldots, n.$$

Now we perform the above calculations for a simple example with the block circulant matrix coming from the quadratic form from the variogram estimator in Chapter 4. Define

$$\begin{aligned} A &= \frac{1}{12}\text{circ}\left(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -1/2 \\ -1/2 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},\right. \\ &\qquad \left.\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -1/2 \\ -1/2 & 0 \end{pmatrix}\right) \end{aligned}$$

Notice that the matrix A of 6×2 by 6×2 ($n = 6, p = 2$) is a block-circulant matrix, we can find P and S_k's, which are given below.

$$P = \frac{1}{\sqrt{6}}(E_1, E_2, E_3, E_4, E_5, E_6)$$

Hence,

$$E_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}, E_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ \frac{1}{2} + i\frac{\sqrt{3}}{2} & 0 \\ 0 & \frac{1}{2} + i\frac{\sqrt{3}}{2} \\ -\frac{1}{2} + i\frac{\sqrt{3}}{2} & 0 \\ 0 & -\frac{1}{2} + i\frac{\sqrt{3}}{2} \\ -1 & 0 \\ 0 & -1 \\ -\frac{1}{2} - i\frac{\sqrt{3}}{2} & 0 \\ 0 & \frac{1}{2} - i\frac{\sqrt{3}}{2} \\ \frac{1}{2} - i\frac{\sqrt{3}}{2} & 0 \\ 0 & \frac{1}{2} - i\frac{\sqrt{3}}{2} \end{pmatrix}, E_3 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -\frac{1}{2} + i\frac{\sqrt{3}}{2} & 0 \\ 0 & -\frac{1}{2} + i\frac{\sqrt{3}}{2} \\ -\frac{1}{2} - i\frac{\sqrt{3}}{2} & 0 \\ 0 & -\frac{1}{2} - i\frac{\sqrt{3}}{2} \\ 1 & 0 \\ 0 & 1 \\ -\frac{1}{2} + i\frac{\sqrt{3}}{2} & 0 \\ 0 & -\frac{1}{2} + i\frac{\sqrt{3}}{2} \\ -\frac{1}{2} - i\frac{\sqrt{3}}{2} & 0 \\ 0 & -\frac{1}{2} - i\frac{\sqrt{3}}{2} \end{pmatrix},$$

$$E_4 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 0 \\ 0 & -1 \\ 1 & 0 \\ 0 & 1 \\ -1 & 0 \\ 0 & -1 \\ 1 & 0 \\ 0 & 1 \\ -1 & 0 \\ 0 & -1 \end{pmatrix}, E_5 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -\frac{1}{2} - i\frac{\sqrt{3}}{2} & 0 \\ 0 & -\frac{1}{2} - i\frac{\sqrt{3}}{2} \\ -\frac{1}{2} + i\frac{\sqrt{3}}{2} & 0 \\ 0 & -\frac{1}{2} + i\frac{\sqrt{3}}{2} \\ 1 & 0 \\ 0 & 1 \\ -\frac{1}{2} - i\frac{\sqrt{3}}{2} & 0 \\ 0 & -\frac{1}{2} - i\frac{\sqrt{3}}{2} \\ -\frac{1}{2} + i\frac{\sqrt{3}}{2} & 0 \\ 0 & -\frac{1}{2} + i\frac{\sqrt{3}}{2} \end{pmatrix},$$

One can verify the following

$$A = P\mathrm{diag}(S_k) * P^*,$$

here P is the unitary matrix, and

$$S_1 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, S_2 = S_6 = \begin{pmatrix} 0 & 0.5 \\ 0.5 & 0 \end{pmatrix}, S_3 = S_5 = \begin{pmatrix} 0 & 1.5 \\ 1.5 & 0 \end{pmatrix}, S_4 = \begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix}$$

For this example, we note that A is also a block circulant symmetric matrix with symmetric block elements. Hence, A can be decomposed by a real orthogonal matrix given below:

$$Q = (Q_1, Q_2, Q_3, Q_4, Q_5, Q_6),$$

such that $A = Q\text{diag}(S_1, S_2, \ldots, S_6)Q^T$, where

$$Q_1 = \frac{1}{\sqrt{6}}\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}, Q_2 = \frac{1}{\sqrt{3}}\begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \\ -1/2 & 0 \\ 0 & -1/2 \\ -1 & 0 \\ 0 & -1 \\ -1/2 & 0 \\ 0 & -1/2 \\ 1/2 & 0 \\ 0 & 1/2 \\ 1 & 0 \\ 0 & 1 \end{pmatrix},$$

$$Q_3 = \frac{1}{\sqrt{3}}\begin{pmatrix} \sqrt{3}/2 & 0 \\ 0 & \sqrt{3}/2 \\ \sqrt{3}/2 & 0 \\ 0 & \sqrt{3}/2 \\ 0 & 0 \\ 0 & 0 \\ -\sqrt{3}/2 & 0 \\ 0 & -\sqrt{3}/2 \\ -\sqrt{3}/2 & 0 \\ 0 & -\sqrt{3}/2 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, Q_4 = \frac{1}{\sqrt{3}}\begin{pmatrix} -1/2 & 0 \\ 0 & -1/2 \\ -1/2 & 0 \\ 0 & -1/2 \\ 1 & 0 \\ 0 & 1 \\ -1/2 & 0 \\ 0 & -1/2 \\ -1/2 & 0 \\ 0 & -1/2 \\ 1 & 0 \\ 0 & 1 \end{pmatrix},$$

$$Q_5 = \frac{1}{\sqrt{3}}\begin{pmatrix} \sqrt{3}/2 & 0 \\ 0 & \sqrt{3}/2 \\ -\sqrt{3}/2 & 0 \\ 0 & -\sqrt{3}/2 \\ 0 & 0 \\ 0 & 0 \\ \sqrt{3}/2 & 0 \\ 0 & \sqrt{3}/2 \\ -\sqrt{3}/2 & 0 \\ 0 & -\sqrt{3}/2 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, Q_6 = \frac{1}{\sqrt{6}}\begin{pmatrix} -1 & 0 \\ 0 & -1 \\ 1 & 0 \\ 0 & 1 \\ -1 & 0 \\ 0 & -1 \\ 1 & 0 \\ 0 & 1 \\ -1 & 0 \\ 0 & -1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

CHAPTER III

COVARIANCE AND VARIOGRAM ESTIMATORS ON A CIRCLE

In this chapter, we first introduce the random process on the circle, followed by stationarity and intrinsic stationarity. In Section 3.2, we present the gridded data structure on the circle as well as the covariance and variogram estimators based on the method of moments (MOM). We investigate some properties of these estimators through spectral decomposition of circulant matrices. Note that part of results in this chapter have been given in Vanlangenberg (2016).

3.1 Random Process on the Circle

Let $\{X(t), t \in [0, 2\pi]\}$ (implicitly we assume $X(0) = X(2\pi)$) be a random process on the circle. Under the assumption of continuity in quadratic means, $X(t)$ can be represented as the following Fourier series (for example, see Dufour and Roy (1976))

$$X(t) = A_0 + \sum_{n=1}^{\infty} A_n \cos(nt) + B_n \sin(nt), \quad t \in [0, 2\pi],$$

where random variables $A_0, A_n, B_n, n = 1, 2, \ldots$ can be obtained by

$$\begin{aligned} A_0 &= \frac{1}{\pi} \int_0^{2\pi} X(t)dt, \\ A_n &= \frac{1}{2\pi} \int_0^{2\pi} X(t) \cos(nt)dt, \\ B_n &= \frac{1}{2\pi} \int_0^{2\pi} X(t) \sin(nt)dt. \end{aligned}$$

The covariance function at two points with angular difference θ is represented by $C(t, \theta) = cov(X(t+\theta), X(t))$, a bivariate function of θ and t.

3.1.1 Stationarity and Intrinsic Stationarity

Assume the underlying process $X(t)$ is *stationary* on the circle, that is, its covariance function solely depends on the (un-directional) angular difference θ.

$$C(\theta) = cov(X(t+\theta), X(t)), \qquad \theta \in [0, \pi],$$

and $E(X(t)) = \mu$ is a possibly unknown constant. Note that under stationarity of $X(t)$, the spectral representation for $C(\theta)$ is given by (for example, Roy (1972))

$$C(\theta) = \sum_{k=0}^{\infty} a_k \cos(k\theta), \quad \theta \in [0, \pi].$$

According to Roy (1972), for example, $X(t)$ is stationary with mean μ if and only if

$$\begin{aligned} E(A_0) &= \mu, E(A_n) = E(B_n) = 0, \\ var(A_0) &= a_0, cov(A_0, A_n) = cov(A_0, B_n) = 0, var(A_n) = var(B_n) = a_n, \\ cov(A_n, A_m) &= cov(B_n, B_m) = 0 (m \neq n), cov(A_n, B_m) = 0 \quad \text{for any } n, m > 0. \end{aligned}$$

Now if we assume that $X(t)$ is intrinsically stationary, that is, $E(X(t)) = \mu$ constant, and the variogram function

$$2\gamma(\theta) = var(X(t+\theta) - X(t)) = E(X(t+\theta) - X(t))^2$$

is solely a function of angular difference θ, then one can show that

$$\gamma(\theta) = C(0) - C(\theta)$$

and $\gamma(\theta)$ has the following spectral representation (Schoenberg, 1942).

$$\gamma(\theta) = \sum_{n=1}^{\infty} a_n(1 - \cos(n\theta)).$$

3.2 Estimation on the Circle

3.2.1 Gridded Data Structure

Note that in real global-scale data, data values are normally observed on grids along each latitude (circle). Hence, throughout this section, we assume that data values observed on the circle are equally spaced with the common interval length $\delta = 2\pi/n$ for n the sample size. For simplicity, we take $n = 2N$, an even number so that $\delta = \pi/N$. All discussions are carried over easily if n is an odd number. Therefore, all points observed on the gridded locations are represented as $\{X(k\delta\}_{k=0}^{n-1}\}$. In general, the angular difference between two points on the circle could take the following sequence of values.

$$0, \pm\delta, \pm 2\delta, \cdots, \pm(N-1)\delta, N\delta = \pi.$$

In the mean time, if $X(t)$ is assumed to be stationary on the circle, we will then have the un-directional angular difference $\Delta\lambda$, that is, $\Delta\lambda = 0, \delta, 2\delta, \cdots, (N-1)\delta, N\delta = \pi$.

3.2.2 Variance-Covariance Matrix of a Gridded Random Vector

Let $\{(X(m\delta),\}_{m=0}^{n-1}$ be the observed random vector on gridded locations of a unit circle. For notational simplicity, we denote $\underset{\sim}{X} = (X_1, X_2, \ldots, X_n)^T$ with $X_j = X((j-1)\delta)$. If $X(t)$ is stationary with constant mean μ, then $E(\underset{\sim}{X}) = \mu\underset{\sim}{1}_n$ and $var(\underset{\sim}{X}) = \Sigma$ given by the following (Note that $C((n-k)2\pi/n) = C(k2\pi/n))$

$$\begin{aligned}\Sigma &= \begin{pmatrix} C(0) & C\left(\frac{2\pi}{n}\right) & C\left(\frac{(N-1)2\pi}{n}\right) & C(\pi) & C\left(\frac{(N-1)2\pi}{n}\right) & \cdots & C\left(\frac{2\pi}{n}\right) \\ C\left(\frac{2\pi}{n}\right) & C(0) & C\left(\frac{(N-2)2\pi}{n}\right) & C\left(\frac{(N-1)2\pi}{n}\right) & C(\pi) & \cdots & C\left(\frac{4\pi}{n}\right) \\ C\left(\frac{4\pi}{n}\right) & C\left(\frac{2\pi}{n}\right) & C\left(\frac{(N-3)2\pi}{n}\right) & C\left(\frac{(N-2)2\pi}{n}\right) & C\left(\frac{(N-1)2\pi}{n}\right) & \cdots & C\left(\frac{6\pi}{n}\right) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ C\left(\frac{2\pi}{n}\right) & C\left(\frac{4\pi}{n}\right) & C(\pi) & C\left(\frac{(N-1)2\pi}{n}\right) & C\left(\frac{(N-2)2\pi}{n}\right) & \cdots & C(0) \end{pmatrix} \\ &= \operatorname{circ}\left(C(0), C(2\pi/n), C(4\pi/n), \cdots, C((N-1)2\pi/n), C(\pi),\right. \\ &\qquad \left. C((N-1)2\pi/n), \cdots, C(2\pi/n)\right).\end{aligned}$$

As a simple example, we consider the case of $n = 6$, we have

$$
\begin{aligned}
\Sigma &= \begin{pmatrix}
C(0) & C(\frac{\pi}{3}) & C(\frac{2\pi}{3}) & C(\pi) & C(\frac{2\pi}{3}) & C(\frac{\pi}{3}) \\
C(\frac{\pi}{3}) & C(0) & C(\frac{\pi}{3}) & C(\frac{2\pi}{3}) & C(\pi) & C(\frac{2\pi}{3}) \\
C(\frac{2\pi}{3}) & C(\frac{\pi}{3}) & C(0) & C(\frac{\pi}{3}) & C(\frac{2\pi}{3}) & C(\pi) \\
C(\pi) & C(2\pi/3) & C(\frac{\pi}{3}) & C(0) & C(\frac{\pi}{3}) & C(\frac{2\pi}{3}) \\
C(\frac{2\pi}{3}) & C(\pi) & C(\frac{2\pi}{3}) & C(\frac{\pi}{3}) & C(0) & C(\frac{\pi}{3}) \\
C(\frac{\pi}{3}) & C(\frac{2\pi}{3}) & C(\pi) & C(\frac{2\pi}{3}) & C(\frac{\pi}{3}) & C(0)
\end{pmatrix} \\
&= circ(C(0), C(\pi/3), C(2\pi/3), C(\pi), C(2\pi/3), C(\pi/3)).
\end{aligned}
$$

From above, we have $E(\underset{\sim}{X} - \mu \underset{\sim}{1}_n) = \underset{\sim}{0}_n$ and $var(\underset{\sim}{X} - \mu \underset{\sim}{1}_n) = \Sigma$. Note that the above Σ is a circulant matrix. Hence it can be diagonalized as

$$
\Sigma = Q \Lambda^{(\Sigma)} Q^T,
$$

where Q is the real orthogonal matrix as given in Section 2.4, and

$$
\Lambda^{(\Sigma)} = \mathrm{diag}(\lambda_1^{(\Sigma)}, \lambda_2^{(\Sigma)}, \cdots, \lambda_n^{(\Sigma)}),
$$

with the eigenvalues given by

$$
\begin{aligned}
\lambda_j^{\Sigma} &= \left(C(0) + 2 \sum_{m=1}^{N-1} C(m\delta) \cos((j-1)m\delta) + C(\pi) \cos((j-1)N\delta) \right) \\
&= \left(C(0) + 2 \sum_{m=1}^{N-1} C(m\delta) \cos((j-1)m\delta) + C(\pi) \cos((j-1)\pi) \right).
\end{aligned}
$$

Throughout the rest of this chapter, we assume that Σ is strictly positive definite, and hence, $\lambda_j^{(\Sigma)} > 0, j = 1, 2, \ldots, n$.

Proposition 3.1. Let $\underset{\sim}{Y} = Q^T(\underset{\sim}{X} - \mu \underset{\sim}{1}_n) := (Y_1, Y_2, \ldots, Y_n)^T$. Then

(1) $E(\underset{\sim}{Y}) = \underset{\sim}{0}, var(\underset{\sim}{Y}) = \Lambda^{(\Sigma)}$, that is, $Y_1, Y_2, \ldots, Y_n$ are uncorrelated with mean zero and variance $var(Y_j) = \lambda_j^{(\Sigma)}$.

(2) If $X(t)$ is a stationary Gaussian random process on the circle, then $\underset{\sim}{Y} \sim N(\underset{\sim}{0}, \Lambda^{(\Sigma)})$, that is, $Y_1, Y_2, \ldots, Y_n$ are independent $Y_j \sim N(0, \lambda_j^{(\Sigma)})$.

Proof. Let $\underset{\sim}{Y} = Q^T \left(\underset{\sim}{X} - \underset{\sim}{1}_n \mu\right)$. Then obviously its mean equals $\underset{\sim}{\mathbf{0}}$ and variance-covariance matrix is given by

$$\begin{aligned} var(\underset{\sim}{Y}) &= cov(Q^T \left(\underset{\sim}{X} - \underset{\sim}{1}_n \mu\right), Q^T \left(\underset{\sim}{X} - \underset{\sim}{1}_n \mu\right)) \\ &= Q^T \Sigma Q = Q^T Q \Lambda^{(\Sigma)} Q^T Q = \Lambda^{(\Sigma)}. \end{aligned}$$

That is, $Y_1, Y_2, \cdots, Y_n$ are uncorrelated random variates, where Y_j has mean zero and variance $\lambda_j^{(\Sigma)}$. The second conclusion follows easily.

Remark 3.1. From Proposition 3.1, without loss of generality, we may assume that $X(t)$ is a zero-mean random process on the circle throughout the rest of this section. Consequently, the observed random vector $\underset{\sim}{X}$ will be assumed to have $E(\underset{\sim}{X}) = \underset{\sim}{0}$ throughout.

3.2.3 Covariance Estimation on the Circle

Now we consider the covariance estimator based on MOM. Recall that $\underset{\sim}{X} = (X_1, X_2, \ldots, X_n)^T$ with $X_j = X((j-1)\delta)$ is a collection of gridded observations on the circle. One obvious unbiased estimator of possibly unknown constant mean μ

is given by $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$. Denote $\Delta\lambda = k\delta$ for $k = 0, 1, 2, \ldots, N = n/2$ (recall $n = 2N$ assumed). The MOM estimator of $C(\Delta\lambda) = C(k\delta)$ is then given by

$$\begin{aligned}\hat{C}(\Delta\lambda) &= \frac{1}{n}\sum_{i=1}^{n}(X((i-1)\delta + \Delta\lambda) - \bar{X})(X((i-1)\delta) - \bar{X}) \\ &= \frac{1}{n}\sum_{i=1}^{n} X((i-1)\delta + \Delta\lambda)X((i-1)\delta) - (\bar{X})^2.\end{aligned}$$

Now we introduce the following rotation matrix R_n, which transforms $\underset{\sim}{X} = (X_1, X_2, \ldots, X_n)^T$ to $\underset{\sim}{X}_{(k+1)} = (X_{k+1}, \ldots, X_n, X_1, \ldots, X_{k-1}, X_k)^T$, that is, $\underset{\sim}{X}_{(k+1)} = R_n\underset{\sim}{X}$, where R_n is given by

$$R_n^T = \begin{pmatrix} 0 & 0 & \ldots & 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & \ldots & 0 & 0 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & 0 & 0 & 0 & \ldots & 1 & 0 \\ 0 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 & 1 \\ 1 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & 0 & 0 & \ldots & 0 & 0 \end{pmatrix}. \qquad (3.1)$$

Here the number 1 in the first row of R_n^T appears at the $(k+1)$th position. Note that $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i = \frac{1}{n}\underset{\sim}{1}_n^T\underset{\sim}{X}$. We then express the covariance estimator as the following quadratic form,

$$
\begin{aligned}
\hat{C}(\Delta\lambda) = \hat{C}(k\delta) &= \frac{1}{n}\underset{\sim}{X}^T_{(k+1)}\underset{\sim}{X} - \left(\frac{1}{n}\underset{\sim}{1}^T_n\underset{\sim}{X}\right)^T * \left(\frac{1}{n}\underset{\sim}{1}^T_n\underset{\sim}{X}\right) \\
&= \frac{1}{n}\underset{\sim}{X}^T R_n^T\underset{\sim}{X} - \left(\frac{1}{n}\underset{\sim}{1}^T_n\underset{\sim}{X}\right)^T * \left(\frac{1}{n}\underset{\sim}{1}^T_n\underset{\sim}{X}\right) \\
&= \frac{1}{n}\underset{\sim}{X}^T\left(R_n^T - \frac{1}{n}\underset{\sim}{1}_n\underset{\sim}{1}^T_n\right)\underset{\sim}{X} \\
&= \underset{\sim}{X}^T M(k\delta)\underset{\sim}{X}.
\end{aligned}
$$

Here

$$
\begin{aligned}
M(k\delta) &= \frac{1}{n}\left(R_n^T - \frac{1}{n}\underset{\sim}{1}_n\underset{\sim}{1}^T_n\right) \\
&= \frac{1}{n}\text{circ}(-1/n, -1/n, \ldots, -1/n, 1-1/n, -1/n, \ldots, -1/n, -1/n).
\end{aligned}
$$

As an example, we consider the case of $n = 6$. We have the following matrices corresponding to $\Delta\lambda = 0, \pi/3, 2\pi/3$, and π, respectively.

$$
\begin{aligned}
M(0) &= \frac{1}{6}\begin{pmatrix}
5/6 & -1/6 & -1/6 & -1/6 & -1/6 & -1/6 \\
-1/6 & 5/6 & -1/6 & -1/6 & -1/6 & -1/6 \\
-1/6 & -1/6 & 5/6 & -1/6 & -1/6 & -1/6 \\
-1/6 & -1/6 & -1/6 & 5/6 & -1/6 & -1/6 \\
-1/6 & -1/6 & -1/6 & -1/6 & 5/6 & -1/6 \\
-1/6 & -1/6 & -1/6 & -1/6 & -1/6 & 5/6
\end{pmatrix} \\
&= \frac{1}{6}circ(5/6, -1/6, -1/6, -1/6, -1/6, -1/6);
\end{aligned}
$$

and

$$
\begin{aligned}
M(\pi/3) &= \frac{1}{6}\begin{pmatrix} -1/6 & 5/6 & -1/6 & -1/6 & -1/6 & -1/6 \\ -1/6 & -1/6 & 5/6 & -1/6 & -1/6 & -1/6 \\ -1/6 & -1/6 & -1/6 & 5/6 & -1/6 & -1/6 \\ -1/6 & -1/6 & -1/6 & -1/6 & 5/6 & -1/6 \\ -1/6 & -1/6 & -1/6 & -1/6 & -1/6 & 5/6 \\ 5/6 & -1/6 & -1/6 & -1/6 & -1/6 & -1/6 \end{pmatrix} \\
&= \frac{1}{6} circ(-1/6, 5/6, -1/6, -1/6, -1/6, -1/6); \\
M(2\pi/3) &= \frac{1}{6}\begin{pmatrix} -1/6 & -1/6 & 5/6 & -1/6 & -1/6 & -1/6 \\ -1/6 & -1/6 & -1/6 & 5/6 & -1/6 & -1/6 \\ -1/6 & -1/6 & -1/6 & -1/6 & 5/6 & -1/6 \\ -1/6 & -1/6 & -1/6 & -1/6 & -1/6 & 5/6 \\ 5/6 & -1/6 & -1/6 & -1/6 & -1/6 & -1/6 \\ -1/6 & 5/6 & -1/6 & -1/6 & -1/6 & -1/6 \end{pmatrix} \\
&= \frac{1}{6} circ(-1/6, -1/6, 5/6, -1/6, -1/6, -1/6) \\
M(\pi) &= \frac{1}{6}\begin{pmatrix} -1/6 & -1/6 & -1/6 & 5/6 & -1/6 & -1/6 \\ -1/6 & -1/6 & -1/6 & -1/6 & 5/6 & -1/6 \\ -1/6 & -1/6 & -1/6 & -1/6 & -1/6 & 5/6 \\ 5/6 & -1/6 & -1/6 & -1/6 & -1/6 & -1/6 \\ -1/6 & 5/6 & -1/6 & -1/6 & -1/6 & -1/6 \\ -1/6 & -1/6 & 5/6 & -1/6 & -1/6 & -1/6 \end{pmatrix} \\
&= \frac{1}{6} circ(-1/6, -1/6, -1/6, 5/6, -1/6, -1/6).
\end{aligned}
$$

Hence $M(k\delta)$ can be diagonalized by a unitary matrix P given in Section 2.4 as following.

$$M(k\delta) = P\mathrm{diag}(\lambda_1^{(M)}, \lambda_2^{(M)}, \ldots, \lambda_n^{(M)})P^*,$$

where $\lambda_j^{(M)}, j = 1, 2, \ldots, n$ are the eigenvalues of $M(k\delta)$.

To compute these eigenvalues, we first decompose $M(k\delta)$ as following.

$$M(k\delta) = -\frac{1}{n^2}\mathrm{circ}(1, 1, \ldots, 1) + \frac{1}{n}\mathrm{circ}(0, 0, \ldots, 0, 1, 0, \ldots, 0) \quad := \quad M_1(k\delta) + M_2(k\delta),$$

where the number 1 in $M_2(k\delta)$ above appears at the $(k+1)$th position. Hence, the eigenvalues for both matrices are given by

$$\lambda_j^{(M_1)} = \begin{cases} -\frac{1}{n}, & j = 1, \\ 0, & j \neq 1, \end{cases} \qquad \lambda_j^{(M_2)} = \frac{1}{n}\omega_j^k.$$

Here $\omega_j^k = \exp(2\pi(j-1)k/n)$. Therefore,

$$\lambda_j^{(M)} \quad = \quad = \begin{cases} \frac{1}{n}\omega_j^k & \text{if } 2 \leq j \leq n, \\ 0 & \text{if } j = 1. \end{cases}$$

Proposition 3.2. The covariance estimator is generally biased. Specifically, for $k = 0, 1, 2, \ldots, N = n/2$,

$$E(\hat{C}(k\delta)) = C(k\delta) - \frac{1}{n}\left(C(0) + 2\sum_{m=1}^{N-1} C(m\delta) + C(\pi)\right).$$

Proof: We first rewrite the eigenvalues of Σ in the complex form, that is,

$$\lambda_j^{(\Sigma)} = \sum_{m=0}^{n-1} C(m\delta)\omega_j^m$$

with $C((n-m)\delta) = C(m\delta)$ for $m = 0, 1, 2, \ldots, N$. Now according to Corollary 1.1 in Chapter 1, we have

$$\begin{aligned}
E(\hat{C}(k\delta)) &= tr(M(k\delta)\Sigma) = \sum_{j=2}^{n} \lambda_j^{(\Sigma)}\lambda_j^{(M)} \\
&= \frac{1}{n}\sum_{j=2}^{n}\omega_j^k \sum_{m=0}^{n-1} C(m\delta)\omega_j^m \\
&= -\frac{1}{n}\left(C(0) + 2\sum_{m=1}^{N-1} C(m\delta) + C(\pi)\right) + \frac{1}{n}\sum_{j=1}^{n}\omega_j^k \sum_{m=0}^{n-1} C(m\delta)\omega_j^m.
\end{aligned}$$

The first term above is added in for $j = 1$ in the original summation. The second term is given by

$$\frac{1}{n}\sum_{j=1}^{n}\omega_j^k \sum_{m=0}^{n-1} C(m\delta)\omega_j^m = \frac{1}{n}\sum_{m=0}^{n-1} C(m\delta)\sum_{j=1}^{n}\omega_j^k\omega_j^m = C(k\delta).$$

The last equality is due to the following.

$$\sum_{j=1}^{n} \omega_j^k \omega_j^m = \begin{cases} n, & \text{if } 0 \leq k = m \leq N, \\ 0, & \text{otherwise.} \end{cases}$$

Our result follows.

Remark 3.2: The result in Proposition 3.2 recovers the one given in Vanlangenberg (2016). Note that one has

$$var(\bar{X}) = \frac{1}{n}\left(C(0) + 2\sum_{m=1}^{N-1} C(m\delta) + C(\pi)\right)$$

leading to $E(\hat{C}(k\delta)) = C(k\delta) - var(\bar{X})$.

3.2.4 Variogram Estimator on the Circle

Now we consider the equal-distance gridded points $\{(i-1) \times 2\pi/n : i = 1, 2, \ldots, n\}$ on the circle. Assume that the random process $X(t)$ is stationary with mean zero. Note that for $\Delta\lambda = k\delta$, with $k = 0, 1, \cdots, N$, the Matheron's classical semivariogram estimator (Matheron, 1962) on the circle can be written as

$$\begin{aligned} \hat{\gamma}(\Delta\lambda) &= \frac{1}{2n}\sum_{i=1}^{n}(X((i-1)\delta + \Delta\lambda) - X((i-1)\delta))^2, \\ &= \frac{1}{2n}(\underset{\sim}{X}_{(k+1)} - \underset{\sim}{X})^T(\underset{\sim}{X}_{(k+1)} - \underset{\sim}{X}), \end{aligned}$$

where $\underset{\sim}{X}_{(k+1)}$ is defined in Section 3.2.3. Hence, we have, for $\Delta\lambda = k\delta, k = 0, 1, \ldots, N$,

$$\hat{\gamma}(\Delta\lambda) = \hat{\gamma}(k\delta) = \frac{1}{2n}\underset{\sim}{X}(R_n^T - I_n)(R_n - I_n)\underset{\sim}{X} = \underset{\sim}{X}A(k\delta)\underset{\sim}{X},$$

where,

$$\begin{aligned}
A(0) &= 0; \\
A(k\delta) &= \frac{1}{2n}circ(2, 0, 0, \ldots, -1, 0, \ldots, -1, 0, \ldots, 0), 1 \leq k \leq N-1, \\
&\quad \text{where } -1\text{'s are placed at } (k+1)^{th} \text{ and } (n-k+1)^{th} \text{ positions;} \\
A(N\delta) &= A(\pi) = \frac{1}{2n}circ(2, 0, 0, \ldots, -2, 0, \ldots, 0), \\
&\qquad \text{where } -2 \text{ is placed at } (N+1)\text{th position.}
\end{aligned}$$

Now we provide an example for computing $A(k\delta)$. Let $n = 6$. We have four angular differences $k\delta = 0, \pi/3, 2\pi/3, \pi$, which gives the following results.

$$
\begin{aligned}
A(0) &= 0 \\
A(\pi/3) &= \frac{1}{6}\begin{pmatrix} 2 & -1 & 0 & 0 & 0 & -1 \\ -1 & 2 & -1 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 \\ 0 & 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & 0 & -1 & 2 & -1 \\ -1 & 0 & 0 & 0 & -1 & 2 \end{pmatrix} = \frac{1}{6}circ(2, -1, 0, 0, 0, -1); \\
A(2\pi/3) &= \frac{1}{6}\begin{pmatrix} 2 & 0 & -1 & 0 & -1 & 0 \\ 0 & 2 & 0 & -1 & 0 & -1 \\ -1 & 0 & 2 & 0 & -1 & 0 \\ 0 & -1 & 0 & 2 & 0 & -1 \\ -1 & 0 & -1 & 0 & 2 & 0 \\ 0 & -1 & 0 & -1 & 0 & 2 \end{pmatrix} = \frac{1}{6}circ(2, 0, -1, 0, -1, 0); \\
A(\pi) &= \frac{1}{6}\begin{pmatrix} 2 & 0 & 0 & -2 & 0 & 0 \\ 0 & 2 & 0 & 0 & -2 & 0 \\ 0 & 0 & 2 & 0 & 0 & -2 \\ -2 & 0 & 0 & 2 & 0 & 0 \\ 0 & -2 & 0 & 0 & 2 & 0 \\ 0 & 0 & -2 & 0 & 0 & 2 \end{pmatrix} = \frac{1}{6}circ(2, 0, 0, -2, 0, 0).
\end{aligned}
$$

Obviously, $A(k\delta)$ is a symmetric circulant matrix. Then $A(k\delta)$ has the following spectral decomposition

$$A(k\delta) = Q\Lambda^{(A)}Q^T,$$

where Q is the real-valued orthogonal Fourier matrix given in Seciton 2.4, and

$$\Lambda^{(A)} = \mathrm{diag}(\lambda_1^{(A)}, \lambda_2^{(A)}, \cdots, \lambda_n^{(A)})$$

with the eigenvalues of $A(k\delta)$ given by

$$\begin{aligned}
\lambda_j^{(A)} &= \frac{1}{2n}(2 - (\exp((j-1)2\pi i/n))^k - (\exp((j-1)2\pi i/n))^{n-k}) \\
&= \frac{1}{2n}(2 - \exp(k(j-1)2\pi i/n) - \exp(-k(j-1)2\pi i/n)) \\
&= \frac{1}{n}(1 - \cos((j-1)k\delta)), \qquad j = 1, 2, \ldots, n
\end{aligned}$$

for $1 \leq k \leq N-1$, and for $k = N$,

$$\begin{aligned}
\lambda_j^{(A)} &= \frac{1}{2n}(2 - 2(\exp((j-1)2\pi i/n))^N) \\
&= \frac{1}{n}(1 - \cos((j-1)\pi)), \qquad j = 1, 2, \cdots, n.
\end{aligned}$$

In summary, the eigenvalues of $A(k\delta)$ are given by $\lambda_j^{(A)} = \frac{1}{n}(1 - \cos(jk\delta)), j = 1, 2, \cdots, n$.

Now we give the following proposition regarding the variogram estimator.

Proposition 3.3. If $X(t)$ is stationary on the circle, the semivariogram estimator $\hat{\gamma}(k\delta), k = 0, 1, \ldots, N$, has the following properties.

(1) $E(\hat{\gamma}(k\delta)) = \gamma(k\delta)$.

(2) $\hat{\gamma}(k\delta) = \sum_{j=1}^{n} \lambda_j^{(A)} Y_j^2$, where $Y_1, Y_2, \cdots Y_n$ are uncorrelated random variates with $E(Y_j) = 0, var(Y_j) = \lambda_j^{(\Sigma)}, j = 1, 2, \ldots, n$.

(3) If $X(t)$ is further assumed to be Gaussian on the circle, then

$$\hat{\gamma}(k\delta) = \sum_{j=1}^{n} \lambda_j^{(A)} \lambda_j^{(\Sigma)} U_j,$$

where $U_1, U_2, \cdots U_n$ are *i.i.d.* χ_1^2 random variates.

Proof: First, from Corollary 1.1, we have

$$\begin{aligned}
E(\hat{\gamma}(k\delta)) &= \frac{1}{n} \sum_{j=1}^{n} \lambda_j^{(A)} \lambda_j^{(\Sigma)} \\
&= \frac{1}{n} \sum_{j=1}^{n} (1 - \cos((j-1)k\delta)) \\
&\quad \left(C(0) + 2 \sum_{k=1}^{N-1} C(k\delta) \cos((j-1)k\delta) + C(\pi) \cos((j-1)N\delta) \right). \\
&= C(0) - \frac{1}{n} \sum_{j=1}^{n} (\cos((j-1)k\delta) \\
&\quad * (2 \sum_{m=1}^{N-1} C(m\delta) \cos((j-1)m\delta) + C(\pi) \cos((j-1)N\delta)) \Bigg) \\
&= \begin{cases} C(0) - \frac{2}{n} C(k\delta) \sum_{j=1}^{n} \cos^2((j-1)k\delta), & \text{if } 0 < k < N-1 \\ C(0) - \frac{1}{n} C(k\delta) \sum_{j=1}^{n} \cos^2((j-1)N\delta), & \text{if } k = N \end{cases} \\
&= C(0) - C(k\delta) = \gamma(k\delta).
\end{aligned}$$

For part (2), we have, from Proposition 3.1, the variogram estimator is then given by

$$\hat{\gamma}(\Delta\lambda) = \underset{\sim}{Y}\Lambda^{(A)}\underset{\sim}{Y} = \sum_{m=1}^{n} \lambda_m^{(A)} Y_m^2.$$

The second result follows from Proposition 3.1 with $\underset{\sim}{Y} = (Y_1, Y_2, \ldots, Y_n)^T$.

For part (3), $\underset{\sim}{Y}$ then follows a multivariate normal $N(\underset{\sim}{0}, \Lambda^{(\Sigma)})$, that is, $Y_1, Y_2, \ldots, Y_n$ are *i.i.d.* normally distributed with $E(Y_j) = 0, var(Y_j) = \lambda^{(\Sigma)}$. Hence we have

$$\begin{aligned}
\hat{\gamma}(\Delta\lambda) &= \underset{\sim}{Y}\Lambda^{(A)}\underset{\sim}{Y} = \sum_{m=1}^{n} \lambda_m^{(A)} Y_m^2 \\
&= \sum_{m=1}^{n} \lambda_m^{(A)} \lambda_m^{(\Sigma)} \frac{Y_m^2}{\sqrt{\lambda_m^{(\Sigma)}}} \\
&= \sum_{m=1}^{n} \lambda_m^{(A)} \lambda_m^{(\Sigma)} U_m.
\end{aligned}$$

Note $\frac{Y_m}{\sqrt{\lambda_m^{(\Sigma)}}} \sim N(0,1)$, and so $U_m = \frac{Y_m^2}{\lambda_m^{(\Sigma)}} \sim \chi_1^2$. The third part then follows.

3.3 Conclusions

In this section, we provide the circulantly-diagonalization of the variance-covariance matrix of observed gridded data on the circle when the underlying process is stationary. We then investigate the unbiasedness of the covariance and variogram estimators based on MOM. We show that the covariance estimator is biased but the variogram estimator is unbiased. When the process on the circle is also Gaussian, we represent the MOM variogram estimator in terms of a linear combination of independent and identically distributed χ_1^2 random variates.

CHAPTER IV

COVARIANCE AND VARIOGRAM ESTIMATORS ON A SPHERE

In this chapter, we first introduce the gridded data structure that is observed on the sphere. In Section 4.2, we discuss the axially symmetric process on the sphere as well as the covariance structure associated with the gridded data observed. We then consider the covariance and variogram estimators on the sphere as well as their properties. Some conclusions are given at the end in Section 4.3.

4.1 Random Processes on the Sphere

Let $\{X(P), P = (\phi, \lambda) \in \mathbb{S}^2\}$ be the random process on a unit sphere, with ϕ and λ representing the latitude and longitude, respectively. Under the assumption of continuity in quadratic means, $X(P)$ can be represented as the following spherical Fourier series (for example, Yaglom (1987))

$$X(P) = X(\phi, \lambda) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} Y_{m,n} e^{im\lambda} \bar{P}_n^m(\cos\phi), \tag{4.1}$$

where $\bar{P}$, the conjugate of P, is the normalized associated Legendre functions with zonal wave number m (normalized in the sense that the squared integral on $[-1, 1]$ is 1), and $Y_{m,n}$ are the coefficients satisfying

$$Y_{m,n} = \int_{\mathbb{S}^2} X(P) e^{-im\lambda} \bar{P}_n^m(\cos\phi) dP.$$

Also notice that if $X(P)$ is real-valued, it can be easily derived that $Y_{m,n} = \bar{Y}_{m,n}$. Both m and n in $Y_{m,n}$ are integers and m gives the longitudinal frequency while $n-m$ corresponds to latitudinal frequency.

Without loss of generality, assuming the process has zero mean, the covariance function for $X(P)$ at two locations $P = (\phi_P, \lambda_P)$ and $Q = (\phi_Q, \lambda_Q)$ has the following form:

$$\begin{aligned} R(P,Q) &= \mathrm{E}(X(P)\overline{X(Q)}) \\ &= \sum_{\nu=0}^{\infty}\sum_{\mu=0}^{\infty}\sum_{m=-\nu}^{\nu}\sum_{n=-\mu}^{\mu} \mathrm{E}(Y_{\nu,m}\bar{Y}_{m,n}) e^{im\lambda_P} P_\nu^m(\cos\phi_P) e^{-in\lambda_Q} \bar{P}_\mu^n(\cos\phi_Q). \end{aligned}$$

4.1.1 Stationarity

A process on the sphere is defined to be *weakly stationary, or homogeneous*, if its first two moments are invariant under any rotation on the sphere (rotation-invariant) (for example see Yaglom (1987)). Equivalently, if $X(P)$ is homogeneous on the sphere, $E(X(P))$ is constant and the covariance function of any two locations $p_1 = (\phi_1, \lambda_1)$ and $p_2 = (\phi_2, \lambda_2)$, $cov(X(p_1), X(p_2))$, depends only on their spherical angle $\theta(p_1, p_2) := \theta$, where $\theta \in [0, \pi]$ and

$$cos(\theta) = \sin\phi_1 \sin\phi_2 + \cos\phi_1 \cos\phi_2 \cos(\lambda_1 - \lambda_2).$$

Let $C(\cdot)$ be the covariance function of a homogeneous random process on the unit sphere. Schoenberg (1942) showed that $C(\theta)$ has the following spectral representation: (see also Yaglom (1961))

$$C(\theta) = \sum_{n=1}^{\infty} a_n P_n(\cos\theta), \quad a_n \geq 0, \sum_{n=1}^{\infty} a_n < \infty, \tag{4.2}$$

where $P_n(\cdot)$ are the Legendre polynomials in the form of

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n}(x^2 - 1)^n, |x| \leq 1.$$

Note that if the covariance function has representation (4.2), it can be shown to be positive definite.

4.1.2 Axially Symmetric Processes on the Sphere

The assumption of homogeneity on the sphere requires the random process has constant mean and its covariance function solely depends on the spherical distance. This assumption is often not satisfied in practice. For example, energy and moisture gradients tend to be stronger at the equator than the pole area. Therefore, geophysical processes are most likely to exhibit symmetry on longitude rather than latitude. In this thesis, we focus on the axially symmetric process on the sphere, that is, we assume the process is stationary on each longitude but nonstationary across latitudes (Jones, 1963).

For an *axially symmetric process* $X(P), P \in \mathbb{S}^2$, the *covariance function* $R(P, Q)$ at two locations $P = (\phi_P, \lambda_P)$, $Q = (\phi_Q, \lambda_Q) \in \mathbb{S}^2$ is given by

$$R(\phi_P, \phi_Q, \lambda_P, \lambda_Q) = R(\phi_P, \phi_Q, \lambda_P - \lambda_Q).$$

That is, the covariance function on two locations depends on the (directional) longitudinal difference.

Using spherical harmonics, the covariance function can be expressed as following:

$$\begin{aligned} R(P,Q) &= R(\phi_1,\phi_2,\lambda_P-\lambda_Q) \\ &= \sum_{m=-\infty}^{\infty}\sum_{\nu=|m|}^{\infty}\sum_{\mu=|m|}^{\infty} f_{\nu,\mu,m} e^{im(\lambda_P-\lambda_Q)} P_n^m(\cos\phi_P)\bar{P}_n^m(\cos\phi_Q), \end{aligned} \tag{4.3}$$

where the matrix $F_m(N) = \{f_{\nu,\mu,m}\}_{\nu,\mu=|m|,|m|+1,\cdots,N}$ must be positive definite for all $N \geq |m|$ and $f_{\nu,\mu,m}$ for each fixed integer m (Huang et al, 2012, Stein, 2007). Furthermore, $R(P,Q)$ satisfies the following conditions.

(1) In general, $R(\phi_P,\phi_Q,\lambda_P-\lambda_Q) \neq R(\phi_P,\phi_Q,\lambda_Q-\lambda_P)$ unless $\phi_P=\phi_Q$ under which both locations are at the same latitude.

(2) However, we have $R(\phi_P,\phi_Q,\lambda_P-\lambda_Q) = R(\phi_Q,\phi_P,\lambda_Q-\lambda_P)$.

The *cross variogram function* on the sphere is defined as:

$$2\gamma(\phi_p,\phi_Q,\Delta\lambda) = E\left((X(\phi_P,\lambda+\Delta\lambda)-X(\phi_P,\lambda))(X(\phi_Q,\lambda+\Delta\lambda)-X(\phi_Q,\lambda))\right).$$

When $\phi_P=\phi_Q$, that is, both P and Q locate on the same latitude (circle), consequently, the above expression reduces to the variogram function on the circle as discussed in Chapter 3.

Note that, if μ_P and μ_Q denote the mean of $X(P)$ on the latitudes ϕ_P and ϕ_Q, respectively,

$$\begin{aligned}
\gamma(\phi_p, \phi_Q, \Delta\lambda) &= \frac{1}{2}E\left((X(\phi_P, \lambda+\Delta\lambda) - X(\phi_P, \lambda))(X(\phi_Q, \lambda+\Delta\lambda) - X(\phi_Q, \lambda))\right) \\
&= \frac{1}{2}E\left(((X(\phi_P, \lambda+\Delta\lambda) - \mu_P) - (X(\phi_P, \lambda) - \mu_P))\right. \\
&\quad * \left.((X(\phi_Q, \lambda+\Delta\lambda) - \mu_Q) - (X(\phi_Q, \lambda) - \mu_Q)\right) \\
&= \frac{1}{2}\left(cov(X(\phi_P, \lambda+\Delta\lambda), X(\phi_Q, \lambda+\Delta\lambda)) - cov(X(\phi_P, \lambda+\Delta\lambda), X(\phi_Q, \lambda)\right) \\
&\quad -cov(X(\phi_P, \lambda), X(\phi_Q, \lambda+\Delta\lambda)) + cov(X(\phi_P, \lambda),) \\
&\quad X(\phi_Q, \lambda))) \\
&= \frac{1}{2}\left(R(\phi_P, \phi_Q, 0) - R(\phi_P, \phi_Q, \Delta\lambda) - R(\phi_P, \phi_Q, -\Delta\lambda) + R(\phi_P, \phi_Q, 0)\right) \\
&= R(\phi_P, \phi_Q, 0) - \frac{1}{2}(R(\phi_P, \phi_Q, \Delta\lambda) + R(\phi_P, \phi_Q, -\Delta\lambda)) \\
&= R(\phi_P, \phi_Q, 0) - \frac{1}{2}(R(\phi_P, \phi_Q, \Delta\lambda) + R(\phi_Q, \phi_P, \Delta\lambda)).
\end{aligned}$$

That is,

$$\gamma(\phi_p, \phi_Q, \Delta\lambda) = R(\phi_p, \phi_Q, 0) - \frac{1}{2}(R(\phi_P, \phi_Q, \Delta\lambda) + R(\phi_Q, \phi_P, \Delta\lambda)). \tag{4.4}$$

One special case of the axially symmetric process is the so-called *longitudinally reversible process*, where the covariance function $R(P, Q)$ satisfies

$$R(\phi_P, \phi_Q, \lambda_P - \lambda_Q) = R(\phi_P, \phi_Q, \lambda_Q - \lambda_P). \tag{4.5}$$

This indicates that the covariance function on two locations depends on the (undirectional) longitudinal difference.

4.2 Covariance and Variogram Estimators on the Sphere

4.2.1 Gridded Data Structure

Note that for real global data, data values are always finite and also observed on grids. For example, in the Microwave Sound Unit data that is used in the literature, observations are recorded on 72 latitudes ($n_l = 72$), ranging from $-88.75°$ to $88.75°$ with $2.5°$ apart between adjacent latitudes. On each latitude, there are 144 observations ($n_L = 144$), ranging from $0°$ to $357.5°$, that is, the common interval length on each latitude is $\delta = 2\pi/n_L = 2.5°$. In general, let each location on a unit sphere be represented by its corresponding latitude and longitude (ϕ, λ). We further let the number of latitudes be n_l and the number of longitudes on each latitudes to be $n = n_L = 2N$, an even number for simplicity. Table 1 shows the gridded data structure on the sphere.

Explicitly, on each latitude, all $n = n_L$ data points are evenly spread. Let $\delta = 2\pi/n_L = 2\pi/n$ be the common interval length. Then all observations on two latitudes ϕ_P and ϕ_Q can be expressed as $\{X(\phi_P, k\delta\}_{k=0}^{n-1}\}$ and $\{X(\phi_Q, k\delta\}_{k=0}^{n-1}\}$. Therefore, the longitudinal difference, $\Delta\lambda$, between two points $P = (\phi_P, \lambda_P)$ and $Q = (\phi_Q, \lambda_Q)$ takes the following sequence of values.

$$\Delta\lambda = 0, \pm\delta, \pm 2\delta, \cdots, \pm(N-1)\delta, N\delta = \pi,$$

or more simply, $\Delta\lambda = k\delta, k = 0, \pm 1, \pm 2, \ldots, \pm(N-1), N$.

Table 1. Gridded Data Structure on the Sphere

Latitude	Observed data on each latitude 0	δ	$\dots$	$(n_L - 1)\delta$	Mean
ϕ_1	X_{11}	X_{12}	$\dots$	X_{1n_L}	$\bar{X}_1$
ϕ_2	X_{21}	X_{22}	$\dots$	X_{2n_L}	$\bar{X}_2$
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
ϕ_{n_l}	$X_{n_l 1}$	$X_{n_l 2}$	$\dots$	$X_{n_l n_l}$	$\bar{X}_{n_l}$

Table 2. Gridded Data Structure on Two Latitudes

Latitude	Observed data on each latitude 0	δ	$\dots$	$(n-1)\delta$	mean Mean
ϕ_1	X_{11}	X_{12}	$\dots$	X_{1n}	$\bar{X}_1$
ϕ_2	X_{21}	X_{22}	$\dots$	X_{2n}	$\bar{X}_2$

4.2.2 Variance-Covariance Matrix for Gridded Data

We now focus on covariance function at two locations, consequently, we assume $n_l = 2$ with the number of longitudes $n_L = n$ (again, for simplicity, $n = 2N$, an even number). Therefore, the observed gridded data structure is given in Table 2.

We then form the above random variables in the following order,

$$\underset{\sim}{X} = (X_{11}, X_{21}, X_{12}, X_{22}, \cdots, X_{1n}, X_{2n})^T.$$

The variance and covariance matrix of $\underset{\sim}{X}$ is now in the form of the block circulant matrix given below

$$\Sigma = \begin{pmatrix} C_0 & C_1 & C_2 & \cdots & C_{n-2} & C_{n-1} \\ C_{n-1} & C_0 & C_1 & \cdots & C_{n-3} & C_{n-2} \\ C_{n-2} & C_{n-1} & C_0 & \cdots & C_{n-4} & C_{n-3} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ C_2 & C_3 & C_4 & \cdots & C_0 & C_1 \\ C_1 & C_2 & C_3 & \cdots & C_{n-1} & C_0 \end{pmatrix}, \tag{4.6}$$

where

$$C_i = \begin{pmatrix} cov(X_{11}, X_{1,i+1}) & cov(X_{11}, X_{2,i+1}) \\ cov(X_{21}, X_{1,i+1}) & cov(X_{21}, X_{2,i+1}) \end{pmatrix}, \qquad i = 0, 1, 2, \cdots, n-1.$$

Here C_i can be treated as the variance-covariance matrix between $\underset{\sim}{X}_1 = (X_{11}, X_{21})^T$ and $\underset{\sim}{X}_i = (X_{1i}, X2i)^T$. Since the process is only assumed to be axially symmetric, hence $cov(X_{11}, X_{2,i+1}) = R(\phi_1, \phi_2, i\delta) \neq cov(X_{21}, X_{1,i+1}) = R(\phi_2, \phi_1, i\delta)$. Consequently, C_i may not be symmetric. However, $C_i = C_{n-i}^T$ since for example, $cov(X_{11}, X_{2,i+1}) = R(\phi_1, \phi_2, i\delta) = R(\phi_2, \phi_1, (n-i)\delta) = cov(X_{21}, X_{1,(n-i)})$. Note that Σ is a block circulant matrix, there exists a unitary matrix P as given in Section 2.5, such that Σ can be diagonalized to have a block-diagonal, that is,

$$\Sigma = P\text{diag}(S_1^{(\Sigma)}, S_2^{(\Sigma)}, \cdots, S_n^{(\Sigma)})P^*,$$

with $S_j^{(\Sigma)} = \sum_{m=0}^{n-1} \omega_j^m C_m$. On the other hand, noting the stationarity on each circle for axially symmetric processes, we could assume μ_1 and μ_2 are the possibly unknown means on those two latitudes, respectively. Now letting $\underset{\sim}{\mu} = (\mu_1, \mu_2, \mu_1, \mu_2, \ldots, \mu_1, \mu_2)^T$ if we subtract the mean vector $\underset{\sim}{\mu}$ from $\underset{\sim}{X}$, the resulting vector becomes a zero-mean random vector on the sphere, and for notational simplicity, we still denote this as $\underset{\sim}{X}$. In summary, throughout the rest of this chapter we will assume that $\underset{\sim}{\mu} = \underset{\sim}{0}$, that is, $X(P)$ is a zero-mean axially symmetric process on the sphere.

Now we consider the longitudinally reversal process on the sphere. Under this assumption, the variance-covariance matrix Σ given by (4.6) becomes a block-symmetric block circulant matrix ($C_i = C_{n-i}$) with symmetric blocks C_i ($C_i^T = C_i$). Hence, one can decompose the variance-covariance matrix Σ as following.

$$\Sigma = Q \text{diag}(U_1, U_2, \ldots, U_n) Q^T$$

where Q is a real orthogonal matrix given in Section 2.5 and $U_j, j = 1, 2, \cdots, n$, is a sequence of 2×2 symmetric matrices given by

$$U_j = C_0 + 2 \sum_{m=1}^{N-1} \cos(m(j-1)\delta) C_m + \cos((j-1)\pi) C_N, j = 1, 2, \ldots, n.$$

Now we define the following random vector,

$$\underset{\sim}{Y} = (Y_{11}, Y_{21}, Y_{12}, Y_{22}, \ldots, Y_{1n}, Y_{2n})^T = (\underset{\sim}{Y}_1^T, \underset{\sim}{Y}_2^T, \ldots, \underset{\sim}{Y}_n^T) = Q^T \underset{\sim}{X},$$

with $\underset{\sim}{Y}_j = (Y_{1j}, Y_{2j})^T$. We have the following proposition.

Proposition 4.1. Under the assumption of longitudinal reversibility of $X(P)$ on the sphere,

(1) $E(\underset{\sim}{Y}) = \underset{\sim}{0}, var(\underset{\sim}{Y}) = \text{diag}(U_1, U_2, \cdots, U_n)$. More specifically, $\underset{\sim}{Y}_1, \underset{\sim}{Y}_2, \ldots, \underset{\sim}{Y}_n$ are independent with $var(Y_j) = U_j, j = 1, 2, \ldots, n$.

(2) If we further assume that $X(P)$ is Gaussian, then

$$\underset{\sim}{Y} \sim N(\underset{\sim}{0}, \text{diag}(U_1, U_2, \cdots, U_n)).$$

Proof. Note that $C_j, 0 \leq j \leq n-1$ is symmetric and so is $U_j^{(\Sigma)}, 1 \leq j \leq n$. Hence, it is sufficient to show that $U_j, 1 \leq j \leq n$ is positive definite. For any vector $\underset{\sim}{a} = (a_1, a_2)^T \in \mathbb{R}^2$ and $j = 1, 2, \ldots, n$,

$$v_j = \underset{\sim}{a}^T U_j \underset{\sim}{a} = \underset{\sim}{a}^T C_0 \underset{\sim}{a} + 2 \sum_{m=1}^{N-1} \cos(m(j-1)\delta) \underset{\sim}{a}^T C_m \underset{\sim}{a} + \cos((j-1)\pi) \underset{\sim}{a}^T C_N \underset{\sim}{a}.$$

Note that $\{v_j, j = 1, 2, \cdots, n\}$ are the eigenvalues of the following symmetric circulant matrix D,

$$\begin{aligned} D &= \text{circ}(\underset{\sim}{a}^T C_0 \underset{\sim}{a}, \underset{\sim}{a}^T C_1 \underset{\sim}{a}, \ldots, \underset{\sim}{a}^T C_N \underset{\sim}{a}, \underset{\sim}{a}^T C_{N-1} \underset{\sim}{a}, \ldots, \underset{\sim}{a}^T C_1 \underset{\sim}{a}) \\ &= \begin{pmatrix} \underset{\sim}{a}^T C_0 \underset{\sim}{a} & \underset{\sim}{a}^T C_1 \underset{\sim}{a} & \ldots & \underset{\sim}{a}^T C_N \underset{\sim}{a} & \underset{\sim}{a}^T C_{N-1} \underset{\sim}{a} & \ldots & \underset{\sim}{a}^T C_2 \underset{\sim}{a} \\ \underset{\sim}{a}^T C_1 \underset{\sim}{a} & \underset{\sim}{a}^T C_0 \underset{\sim}{a} & \underset{\sim}{a}^T C_1 \underset{\sim}{a} & \ldots & \underset{\sim}{a}^T C_N \underset{\sim}{a} & \underset{\sim}{a}^T C_{N-1} \underset{\sim}{a} & \ldots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \\ \underset{\sim}{a}^T C_1 \underset{\sim}{a} & \ldots & \underset{\sim}{a}^T C_N \underset{\sim}{a} & \underset{\sim}{a}^T C_{N-1} \underset{\sim}{a} & \ldots & \underset{\sim}{a}^T C_1 \underset{\sim}{a} & \underset{\sim}{a}^T C_0 \underset{\sim}{a} \end{pmatrix}, \end{aligned}$$

which is positive definite due to the positive definiteness of Σ. To see this, for any $\underset{\sim}{b} = (b_1, b_2, \ldots, b_n)^T \in \mathbb{R}^n$, and letting $\underset{\sim}{w} = (b_1\underset{\sim}{a}, b_2\underset{\sim}{a}, \ldots, b_n\underset{\sim}{a})^T$, one can notice that

$$\begin{aligned}
\underset{\sim}{b}^T D \underset{\sim}{b} &= \underset{\sim}{b}^T \begin{pmatrix} \underset{\sim}{a}^T C_0 \underset{\sim}{a} & \underset{\sim}{a}^T C_1 \underset{\sim}{a} & \ldots & \underset{\sim}{a}^T C_N \underset{\sim}{a} & \underset{\sim}{a}^T C_{N-1} \underset{\sim}{a} & \ldots & \underset{\sim}{a}^T C_2 \underset{\sim}{a} \\ \underset{\sim}{a}^T C_1 \underset{\sim}{a} & \underset{\sim}{a}^T C_0 \underset{\sim}{a} & \underset{\sim}{a}^T C_1 \underset{\sim}{a} & \ldots & \underset{\sim}{a}^T C_N \underset{\sim}{a} & \underset{\sim}{a}^T C_{N-1} \underset{\sim}{a} & \ldots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \\ \underset{\sim}{a}^T C_1 \underset{\sim}{a} & \ldots & \underset{\sim}{a}^T C_N \underset{\sim}{a} & \underset{\sim}{a}^T C_{N-1} \underset{\sim}{a} & \ldots & \underset{\sim}{a}^T C_1 \underset{\sim}{a} & \underset{\sim}{a}^T C_0 \underset{\sim}{a} \end{pmatrix} \underset{\sim}{b} \\
&= \underset{\sim}{w}^T \Sigma \underset{\sim}{w} \geq 0,
\end{aligned}$$

by the positive definiteness of Σ, implying the positive definiteness of D. Therefore, $\{v_j, j = 1, 2, \cdots, n\}$ are nonnegative, and so U_j is positive definite. The rest is a simple application from Mathematical Statistics.

4.2.3 Covariance Estimator on the Sphere

Next we consider the cross-covariance estimator for axially symmetric processes. The cross-covariance estimator of two points on latitudes ϕ_1 and ϕ_2 with longitudinal difference $\Delta\lambda = k\delta, k = 0, \pm 1, \pm 2, \ldots, \pm(N-1), N$, is given by

$$\begin{aligned}
\hat{R}_{12}(\Delta\lambda) &= \hat{R}(\phi_1, \phi_2, \Delta\lambda) \\
&= \frac{1}{n} \sum_{i=1}^{n} \big((X(\phi_1, (i-1)\delta + \Delta\lambda) - \bar{X}_{\phi_1})(X(\phi_2, (i-1)\delta) - \bar{X}_{\phi_2}) \big) \\
&= \frac{1}{n} \sum_{i=1}^{n} X(\phi_1, (i-1)\delta + \Delta\lambda) X(\phi_2, (i-1)\delta) - \bar{X}_1 \bar{X}_2.
\end{aligned}$$

The first term can be written as the quadratic form

$$\frac{1}{n}\sum_{i=1}^{n} X(\phi_1,(i-1)\delta+\Delta\lambda)X(\phi_2,(i-1)\delta) = \utilde{X}^T A_1(\Delta\lambda)\utilde{X},$$

where $A(\Delta\lambda)$ takes the following form.

$$A_1(\Delta\lambda) \quad = \quad \frac{1}{n}\text{circ}\left(\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},\ldots,\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},\ldots,\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},\ldots,\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}\right)$$

The second term is expressed as the following quadratic form.

$$\bar{X}_1\bar{X}_2 = \utilde{X}^T A_2(\Delta\lambda)\utilde{X} = \frac{1}{n^2}\utilde{X}^T\text{circ}\left(\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},\ldots,\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},\ldots,\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\right)\utilde{X}$$

such that

$$A_2(\Delta\lambda) = \frac{1}{n^2}\text{circ}\left(\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},\ldots,\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},\ldots,\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\right).$$

Therefore,

$$\hat{R}_{12}(\Delta\lambda) = \utilde{X}^T(A_1(\Delta\lambda) - A_2(\Delta\lambda))\utilde{X} = \utilde{X}^T A(\Delta\lambda)\utilde{X},$$

with

$$A(\Delta\lambda) = \frac{1}{n}\text{circ}\left(\begin{pmatrix} 0 & -1/n \\ 0 & 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 & 1-1/n \\ 0 & 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 & -1/n \\ 0 & 0 \end{pmatrix}\right),$$

which is a block-circulant matrix. According to Section 2.5, there exists a unitary matrix P, such that

$$A(\Delta\lambda) = P\text{diag}(S_1^{(A)}, S_2^{(A)}, \cdots, S_n^{(A)})P^*,$$

with

$$S_j^{(A)} = \sum_{m=0}^{n-1} \omega_j^m A_m,$$

where $A_m, m = 0, 1, \ldots, n-1$, represent the 2×2 blocks in $A(\Delta\lambda)$. With the parallel calculations as given in Section 3.2, one can obtain the following.

$$S_j^{(A)} = \frac{1}{n}\begin{pmatrix} 0 & \omega_j^k \\ 0 & 0 \end{pmatrix} = \frac{1}{n}\omega_j^k \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$

if $2 \leq j \leq n$, and $\mathbf{0} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ if $j = 1$. Then we have the following proposition.

Proposition 4.2. $E(\hat{R}_{12}(k\delta)) = R_{12}(k\delta) - \frac{1}{n}\sum_{m=0}^{n-1} R_{12}(m\delta)$.

Proof: Let $\Delta\lambda = k\delta$. Note that $\hat{R}_{12}$ can be written as the following quadratic form,

$$\hat{R}_{12}(k\delta) = \underset{\sim}{X}^T A(k\delta) \underset{\sim}{X}.$$

We have

$$\begin{aligned}
E[\hat{R}_{12}(k\delta)] &= tr(A(k\delta)\Sigma) \\
&= tr(P\text{diag}(S_1^{(A)}, S_2^{(A)}, \cdots, S_n^{(A)})P^* P\text{diag}(S_1^{(\Sigma)}, S_2^{(\Sigma)}, \cdots, S_n^{(\Sigma)})P^*) \\
&= tr(\sum_{j=1}^{n} S_j^{(A)} * S_j^{(\Sigma)}) \\
&= \frac{1}{n}\sum_{j=2}^{n}\sum_{m=0}^{n-1} \omega_j^k \omega_j^m tr\left(\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} cov(X_{11}, X_{1m}) & cov(X_{11}, X_{2m}) \\ cov(X_{21}, X_{1m}) & cov(X_{21}, X_{2m}) \end{pmatrix}\right) \\
&= \frac{1}{n}\sum_{j=2}^{n}\sum_{m=0}^{n-1} \omega_j^k \omega_j^m R_{12}((m-1)\delta) \\
&= \frac{1}{n}\sum_{m=0}^{n-1} R_{12}(m\delta) \sum_{j=2}^{n} \omega_j^k \omega_j^m \\
&= \frac{1}{n}\sum_{m=0}^{n-1} R_{12}(m\delta) \sum_{j=1}^{n} \omega_j^k \omega_j^m - \frac{1}{n}\sum_{m=0}^{n-1} R_{12}(m\delta) \\
&= R_{12}(k\delta) - \frac{1}{n}\sum_{m=0}^{n-1} R_{12}(m\delta).
\end{aligned}$$

Our result follows.

Remark 4.1. Note that

$$cov(\bar{X}_1, \bar{X}_2) = \frac{1}{n}\sum_{m=0}^{n-1} R_{12}(m\delta),$$

hence,

$$E[\hat{R}_{12}(\Delta\lambda)] = R_{12}(\Delta\lambda) - cov(\bar{X}_1, \bar{X}_2),$$

which implies that the covariance estimator is generally biased.

4.2.4 Variogram Estimator on the Sphere

Next we consider the cross-variogram estimator for axially symmetric processes on the sphere. The cross (semi-)variogram estimator based on MOM is defined as following.

$$\begin{aligned}
\hat{\gamma}_{12}(\Delta\lambda) &= \hat{\gamma}(\phi_1, \phi_2, \Delta\lambda) \\
&= \frac{1}{2n}\sum_{i=1}^{n} ((X(\phi_1, (i-1)\delta + \Delta\lambda) - X(\phi_1, (i-1)\delta)) \\
&\quad (X(\phi_2, (i-1)\delta + \Delta\lambda) - X(\phi_2, (i-1)\delta)))
\end{aligned}$$

If we write the cross-variogram estimator into a quadratic form,

$$\hat{\gamma}_{12}(\Delta\lambda) = \underset{\sim}{X}^T A(\Delta\lambda) \underset{\sim}{X},$$

then $A(\Delta\lambda)$ is a block-symmetric block circulant matrix with symmetric blocks. As a simple example, we consider the case of $n = 6$. Here we only list the matrices when $\Delta\lambda = 0, \pi/3, 2\pi/3$, respectively.

$$
\begin{aligned}
A(0) &= \mathbf{0}_{12\times 12} \\
A(\pi/3) &= \frac{1}{12}\text{circ}\left(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -\frac{1}{2} \\ -\frac{1}{2} & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \right. \\
&\left. \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -\frac{1}{2} \\ -\frac{1}{2} & 0 \end{pmatrix}\right) \\
A(2\pi/3) &= \frac{1}{12}\text{circ}\left(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -\frac{1}{2} \\ -\frac{1}{2} & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \right. \\
&\left. \begin{pmatrix} 0 & -\frac{1}{2} \\ -\frac{1}{2} & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}\right), \\
A(\pi) &= \frac{1}{12}\text{circ}\left(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}, \right. \\
&\left. \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}\right).
\end{aligned}
$$

From Section 2.5, there exists a unitary matrix P such that, for $k = 0, \pm 1, \pm 2, \ldots, \pm(N-1), N$,

$$
A(k\delta) = P * \text{diag}\left(S_1^{(A)}, S_2^{(A)}, \ldots, S_n^{(A)}\right) * P^*
$$

with

$$S_j^{(A)} = (1-\cos((j-1)k\delta))\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad j = 1, 2, \cdots, n.$$

Hence,

$$A(k\delta) = P * \text{diag}\left((1-\cos((j-1)k\delta))\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\right) * P^*.$$

Now we consider the unbiasedness of the cross-variogram estimator.

Proposition 4.3. $E(\hat{\gamma}(k\delta)) = \gamma(k\delta), k = 0, \pm1, \pm2, \ldots, \pm(N-1), N$. *Proof.* From Corollary 1.1,

$$\begin{aligned}
E(\hat{\gamma}_{12}(k\delta)) &= \frac{1}{2n} tr(\sum_{j=1}^{n} S_j^{(A)} * S_j^{(\Sigma)}) \\
&= \frac{1}{2n}\sum_{j=1}^{n}\sum_{m=1}^{n}(1-\cos((j-1)k\delta))\omega^{(j-1)(m-1)} \\
&\quad \times tr\left(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} cov(X_{11}, X_{1m}) & cov(X_{11}, X_{2m}) \\ cov(X_{21}, X_{1m}) & cov(X_{21}, X_{2m}) \end{pmatrix}\right) \\
&= \frac{1}{2n}\sum_{j=1}^{n}\sum_{m=1}^{n}(1-\cos((j-1)k\delta)) \\
&\quad \omega_j^{(m-1)}(R_{21}((m-1)\delta) + R_{12}((m-1)\delta)) \\
&= \frac{1}{2n}\sum_{m=0}^{n-1}(R_{21}(m\delta) + R_{12}(m\delta)) \\
&\quad \sum_{j=0}^{n-1}(1-\cos(jk\delta))(\cos(jm\delta) + i\sin(jm\delta)).
\end{aligned}$$

Now

$$\begin{aligned}
\sum_{j=0}^{n-1}(\cos(jm\delta)+i\sin(jm\delta)) &= \begin{cases} n, & \text{if } m=0; \\ 0, & \text{otherwise.} \end{cases} \\
\sum_{j=0}^{n-1}\cos(jk\delta)\cos(jm\delta) &= \sum_{j=0}^{n-1}\frac{1}{2}(\cos(j(k-m)\delta)+\cos(j(k+m)\delta)) \\
&= \begin{cases} \frac{n}{2}, & \text{if } m=k \text{ or } k+m=n; \\ 0, & \text{otherwise.} \end{cases} \\
\sum_{j=0}^{n-1}\cos(jk\delta)\sin(jm\delta) &= \sum_{j=0}^{n-1}\frac{1}{2}(\sin(j(m-k)\delta)+\sin(j(m+k)\delta)) = 0.
\end{aligned}$$

Therefore, we have

$$\begin{aligned}
E(\hat{\gamma}_{12}(k\delta)) &= \frac{1}{2}((R_{21}(0)+R_{12}(0)) - \frac{1}{4}(R_{21}(k\delta)+R_{12}(k\delta)) \\
&\quad -\frac{1}{4}(R_{21}((n-k)\delta)+R_{12}((n-k)\delta)) \\
&= R_{12}(0)-\frac{1}{2}(R_{21}(k\delta)+R_{12}(k\delta)) = \gamma(k\delta),
\end{aligned}$$

from (4.4). Here we note that

$$R_{12}((n-k)\delta) = R_{21}(k\delta), R_{21}((n-k)\delta) = R_{12}(k\delta), k = 0, \pm 1, \pm 2, \ldots, \pm(N-1), N.$$

Now we assume that $X(P)$ is a longitudinally reversible process on the sphere. We apply Proposition 4.1 to obtain the following result.

Proposition 4.4. Under the assumption of longitudinal reversibility of $X(P)$, then we have

$$\hat{\gamma}(k\delta) = \frac{1}{n}\sum_{m=1}^{n}(1-\cos((m-1)k\delta))Y_{1m}Y_{2m}, \qquad k = 0, 1, \ldots, N,$$

where $\underset{\sim}{Y}_m = (Y_{1m}, Y_{2m})^T, m = 1, 2, \ldots, n,$ are uncorrelated. If $X(P)$ is further assumed to be Gaussian, $\underset{\sim}{Y}_m = (Y_{1m}, Y_{2m})^T, m = 1, 2, \ldots, n$ are independent, each following $N(\underset{\sim}{0}, U_m)$ with U_m given in Proposition 4.1.

Proof. Under the assumption of longitudinal reversibility for $X(P)$, $A(\Delta\lambda)$ can be decomposed, through a real orthogonal matrix Q given in Section 2.5, as a block-diagonal matrix.

$$\begin{aligned}
A(k\delta) &= Q\mathrm{diag}(S_1^{(A)}, S_2^{(A)}, \ldots, S_n^{(A)})Q^T, \qquad \text{hence,} \\
\hat{\gamma}(k\delta) &= \underset{\sim}{Y}^T\mathrm{diag}(S_1^{(A)}, S_2^{(A)}, \ldots, S_n^{(A)})\underset{\sim}{Y} \\
&= \sum_{m=1}^{n} \underset{\sim}{Y}_m^T S_m^{(A)} \underset{\sim}{Y}_m \\
&= \frac{1}{2n}\sum_{m=1}^{n}(1-\cos((m-1)k\delta))(Y_{1m}Y_{2m} + Y_{2m}Y_{1m}) \\
&= \frac{1}{n}\sum_{m=1}^{n}(1-\cos((m-1)k\delta))Y_{1m}Y_{2m}.
\end{aligned}$$

The first part now follows from Proposition 4.1, so is the second part.

4.3 Conclusions

In this chapter, we establish the block circulant variance-covariance structure for the observed gridded random vector on the sphere when the underlying process on the sphere is axially symmetric. We investigate the unbiasedness of the covariance and variogram estimators based on the method of moments. We show that the covariance estimator is biased while the variogram estimator is unbiased. In particular, under the assumption of longitudinal reversibility for the process, the variance-covariance matrix of the random vector can be spectrally decomposed as a block diagonal matrix through a real-valued orthogonal matrix. Consequently, the MOM variogram estimator can be represented as a linear combination of independent random variates if the process is further assumed to be Gaussian.

CHAPTER V

FUTURE RESEARCH

We can extend this thesis work to a number of future research areas. First, we will calculate the variance of the covariance and variogram estimators on both circles and spheres. From Corollary 1.1, if we assume the process is Gaussian, we can then have a closed form for the variance (variogram estimator, for example) as following.

$$var(\hat{\gamma}(k\delta)) = 2tr(A(k\delta)\Sigma A(k\delta)\Sigma).$$

However, the simplification of the above expression is tedious and complicated, which needs further exploration. The derivation of the above variance under the spherical setting might help us establish the possible non-consistency of the MOM variogram estimator, as its non-consistency has been established on the circle in Vanlangenberge (2016). In addition, when the underlying process on the sphere is axially symmetric, the block-diagonalization of the variance-covariance matrix of the observed random vector involves the possible complex random vectors, which provides another research area for further investigation.

REFERENCES

[1] Cressie, N. (1993). *Statistics for Spatial Data*, revised ed. Wiley, New York.

[2] Dufour, J.-M. and Roy, R. (1976). On spectral estimation for a homogeneous random process on the circle. *Stochastic Processes and their Applications.* **4**, 107-120.

[3] Huang, C., Zhang, H., and Robeson, S. (2012). A simplified representation of the covariance structure of axially symmetric processes on the sphere. *Statistics and Probability Letters*, **82**, 1346-1351.

[4] Jones, A. H. (1963). Stochastic processes on a sphere. *Annals of Mathematical Statistics,* **34**, 213 - 217.

[5] Matheron, G. (1962). Traite de Geostatistique Appliquee, Tome I. *Memoires du Bureau de Recherches Geologiques et Minieres*, **14**, Editions Bureau de Recherche Geologiques et Minieres, Paris.

[6] Roy, R. (1972). Spectral analysis for a random proces on the circle. *Journal of Applied Probability*, **9**, 745 - 757.

[7] Schoenberg, I.J. (1942). Positive definite functions on spheres. *Duke Mathematics Journal*, **9**, 96-108.

[8] Seber, G. and Lee, J. (2003). *Linear Regression Analysis.* 2nd ed., Wiley, New Jersey.

[9] De Mazancourt, T and Gerlic, D. (1983). The inverse of a block-circulant matrix. *IEEE Transactions on antennas and propagation*, **AP-31**, 808-810.

[10] Stein, M. (2007). Spatial variation of total column ozone on a global scale. *Annals of Applied Statistics,* **1**, 191-210.

[11] Vanlangenberg, C. (2016). *Data Generation and Estimation for Axially Symmetric Processes on the Sphere.* Ph.D. dissertation. University of North Carolina at Greensboro.

[12] Yaglom, A. M. (1961). Second-order homogeneous random fields. Fourth Berkeley Symposium on Mathematical Statistics and Probability, **2**, 593 - 622, Berkeley, University of California Press.

Yaglom, A. M. (1987). *Correlation Theory of Stationary and Related Random Functions*, Vol. I, Springer, New York.